PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : A61K 35/14, C07K 1/00, 14/00, C07H 21/04, A01N 43/04 — A1

(11) International Publication Number: **WO 00/54787**

(43) International Publication Date: 21 September 2000 (21.09.00)

(21) International Application Number: PCT/US00/06934

(22) International Filing Date: 16 March 2000 (16.03.00)

(30) Priority Data:
60/124,609 16 March 1999 (16.03.99) US

(71) Applicants: THE CHILDREN'S HOSPITAL OF PHILADELPHIA [US/US]; Abramson Pediatric Research Center, 34th & Civic Center Boulevard, Philadelphia, PA 19104–4318 (US). UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL [US/US]; Chapel Hill, NC 27599–3280 (US).

(72) Inventors: HIGH, Katherine, A.; 201 Greenway Lane, Merion, PA 19066 (US). CAMIRE, Rodney, M.; 241–2 Lucas Lane, Voorhees, NJ 08043 (US). LARSON, Peter, J.; 1099 Mandana Boulevard, Oakland, CA 94610 (US). STAFFORD, Darrel, W.; 300 Rainbow Drive, Canboro, NC 27599 (US).

(74) Agents: DOYLE, Kathryn et al.; Akin, Gump, Strauss, Hauer & Feld, L.L.P., One Commerce Square, Suite 2200, 2005 Market Street, Philadelphia, PA 19103–7086 (US).

(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: ENHANCED GAMMA–CARBOXYLATION OF RECOMBINANT VITAMIN K–DEPENDENT CLOTTING FACTORS

(57) **Abstract**

The invention relates to methods of optimizing gamma carboxylation of a vitamin K–dependent protein, methods of generating fully gamma carboxylated vitamin K–dependent protein, and compositions comprising chimeric nucleic acids and proteins for use in treatment of vitamin K–dependent disease states.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

TITLE OF THE INVENTION

ENHANCED GAMMA-CARBOXYLATION OF RECOMBINANT VITAMIN K- DEPENDENT CLOTTING FACTORS

BACKGROUND OF THE INVENTION

Biologically active human blood clotting Factor X is fully gamma carboxylated. Factor X, a vitamin K-dependent two-chain glycoprotein, is a substrate for both the extrinsic (tissue factor/FVIIa) and intrinsic (FVIIIa/FIXa) tenase complexes thus linking these two pathways (Kalafatis et al.,1994, Biochem. Biophys. Acta 1227:113). The activated form of Factor X (FXa) is the serine protease component of the enzymatic complex termed prothrombinase, the only known physiological activator of prothrombin. Prothrombinase assembles through reversible interactions between FXa and the cofactor factor Va (FVa) on an appropriate membrane (i.e., platelet) surface in the presence of Ca^{2+} ions (Mann et al., 990, Blood 76:1). While FXa catalyzes prothrombin activation, the macromolecular interactions which stabilize prothrombinase lead to a substantial enhancement in catalytic efficiency (Mann et al.,1988, Ann. Rev. Biochem. 57:915), indicating that assembly of this complex is an important requisite for rapid and localized thrombin generation. Because Factor X/FXa occupies a central position in the coagulation pathway, there is considerable interest in its therapeutic modulation (Hauptmann et al.,1999, Thromb. Res. 93:203), highlighting the need to better understand structural determinants on Factor X/FXa important to its function.

While extensive progress has been made in delineating structural determinants important for function on thrombin, FIXa, FVIIa, and activated protein C (APC), less is known about FXa. One explanation is the limited number of naturally

occurring FXa mutations to study. Another reason is the difficulty in producing (as compared to other vitamin K-dependent proteins) functional recombinant Factor X/FXa (rFXa). As with all vitamin K-dependent proteins, the biosynthesis of Factor X is complex, involving several co- and post-translational modifications (Kaufman RJ, 1998, Thromb. Haemost. 79:1068). Efficient processing and release of mature two-chain Factor X into the circulation requires, 1) removal of the signal sequence, 2) formation of disulfide bonds, 3) modification of amino-terminal glutamic acid residues to γ-carboxyglutamic acid, 4) modification of one aspartic acid in the first epidermal growth factor (EGF) domain to β-hydroxyaspartic acid, 5) addition of N- and O-linked oligosaccharides to the activation peptide, 6) removal of an internal tripeptide to yield two chain Factor X, and 7) removal of the propeptide just prior to secretion (for review see Kaufman RJ, 1998, Thromb. Haemost. 79:1068). While some of these modifications do not appear essential for Factor X function, the removal of the signal sequence, propeptide, internal tripeptide, and full γ-carboxylation are all steps which are important requisites for the production of biologically active Factor X/FXa.

Expression of rFactor X is heterogeneous with respect to removal of the internal tripeptide, propeptide cleavage, and γ-carboxylation. Expression of rFactor X/FXa in CHO and COS-1 cells appears less efficient than HEK 293 cells with respect to these modifications (Messier et al., 1991, Gene 99:291; Wolf et al., 1991, J. Biol. Chem. 266:13726; Rudolph et al., 1997, Protein Expression and Purification 10:373; Sinha et al., 1994, Thromb. Res. 75:427; Larson et al., 1998, Biochemistry 37:5029). Some of these inefficient modifications can be overcome by expressing rFactor X in HEK 293 cells, cotransfecting with PACE/furin, and modifying the Factor X propeptide at position –2 (Thr→Arg; henceforth referred to as native rwtFactor X). However, inefficient γ-carboxylation still remains a major problem (Rudolph et al., 1997, Protein Expression and Purification 10:373; Larson et al., 1998, Biochemistry 37:5029). For example, it has been discovered that on average only 32% of the rFactor X produced by HEK 293 cells is fully γ-carboxylated while the remaining material

exhibits no γ-carboxylation (Larson et al., 1998, Biochemistry 37:5029). While separation of uncarboxylated and fully γ-carboxylated rFactor X can be readily accomplished, the resulting protein yields are less than desirable. This heterogeneity in γ-carboxylation can be overcome completely by expressing Gla-domainless rFactor X (Rezaie et al., 1993, J Biol. Chem. 268:8176); however, this is a less than satisfactory solution for studies involving macromolecular complex assembly of Factor X/FXa which requires a membrane surface. Thus, an ideal expression system would direct high-level protein production (>2-5 μg rFactor X/10^6 cells/24 hour) while still allowing for efficient execution of post-translational modifications essential to Factor X/FXa function.

The enzyme responsible for modification of glutamic acid residues to γ-carboxyglutamic acid (Gla) in the amino-terminal portion of a number of blood coagulation proteins is the vitamin K-dependent γ-glutamyl carboxylase (Wright et al., 1995, Vitamin K-Dependent g-Glutamyl Carboxylase, in High KA, Roberts HR (eds): Molecular Basis of Thrombosis and Hemostasis, New York, Marcel Dekker, Inc., p 309). The mechanism by which the carboxylase recognizes its substrate is believed to be through initial binding to an 18 amino acid propeptide sequence on the vitamin K-dependent protein (for review see Furie et al., 1990, Blood 75:1753). The importance of the propeptide sequence for γ-carboxylation is demonstrated by studies which show that disruption of this site in FIX, protein C, or prothrombin yield a mature protein that either lacks or is deficient in γ-carboxylation (Jorgensen et al., 1987, Cell 48:185; Foster et al., 1987, Biochemistry 26:7003, Furie et al., 1990, Blood 75:1753), indicating that the propeptide is required for γ-carboxylation. Analysis of naturally occurring mutations in this region supports this conclusion (Chu et al., 1996, J. Clin. Invest. 98:1619; Stanley et al., 1999, Biochemistry 38:15681). Recent studies also support the notion that the γ-carboxylation recognition site on the propeptide is sufficient to direct γ-carboxylation of glutamic acid residues as long as these residues are within 40 amino acids of the γ-carboxylation recognition site (Furie

et al., 1997, J. Biol. Chem. 272:28258).

As noted above, in order that Factor X is biologically active, it must be fully gamma-carboxylated. Until the present invention, it has only been possible to produce biologically active rFactor X which is about 20-40% gamma carboxylated (Larson et al., 1998, Biochemistry 37:5029-5038). There is thus a great need in the art for methods of producing rFactor X which is fully carboxylated. In addition, there is also a great need for the development of methods of producing other mature vitamin K-dependent proteins that are fully gamma carboxylated. The present invention satisfies these needs.

SUMMARY OF THE INVENTION

The invention relates to an isolated chimeric nucleic acid comprising a nucleic acid sequence encoding a propeptide fused to a nucleic acid sequence encoding a vitamin K-dependent protein.

In one aspect, the vitamin K-dependent protein is selected from the group consisting of Factor X, Factor VII, protein S, Factor IX, protein C and prothrombin.

In another aspect, the propeptide is selected from the group consisting of altered or unaltered Factor X, Factor VII, protein S, Factor IX, protein C and prothrombin propeptide.

In other aspects, there is included a vector and a cell comprising the chimeric nucleic acid of the invention.

The invention additionally includes a chimeric protein comprising a propeptide fused to a vitamin K-dependent protein.

In one embodiment, the vitamin K-dependent protein is selected from the group consisting of Factor X, Factor VII, protein S, Factor IX, protein C and prothrombin. In another embodiment, the propeptide is selected from the group

consisting of altered or unaltered Factor X, Factor VII, protein S, Factor IX, protein C and prothrombin propeptide.

The invention further includes a cell comprising the chimeric protein of the invention.

In addition, there is included a method of optimizing the gamma carboxylation of a vitamin K-dependent protein. The method comprises introducing into a cell the chimeric nucleic acid of the invention, expressing the chimeric nucleic in the cell, and assessing the level of gamma carboxylation of the vitamin K-dependent protein expressed by the chimeric nucleic acid, wherein the effect of the propeptide sequence on the gamma carboxylation is measured, the method further comprising modifying the nucleic acid encoding the propeptide until optimal gamma carboxylation of the vitamin K-dependent protein is achieved.

In one embodiment, the modifying includes substituting nucleic acid encoding the propeptide sequence with an altered or different propeptide sequence.

Also included is a method of producing a fully gamma carboxylated vitamin K-dependent protein. The method comprises introducing into a cell an isolated chimeric nucleic acid comprising a nucleic acid encoding a propeptide fused to a nucleic acid sequence encoding a vitamin K-dependent protein and expressing the protein therefrom, thereby producing a fully gamma carboxylated vitamin K-dependent protein.

There is also included a fully gamma carboxylated vitamin K-dependent protein made by the aforementioned method.

Further included is a method of alleviating a vitamin K-dependent protein associated disease in a mammal. This method comprises administering a fully gamma carboxylated protein to a mammal having the disease thereby alleviating the disease.

In addition, there is included a method of alleviating a vitamin K-dependent protein associated disease in a mammal. The method comprises administering the isolated chimeric nucleic acid of the invention to a mammal having

the disease, wherein the chimeric nucleic acid is expressed in a cell in the mammal to produce a fully gamma carboxylated vitamin K-dependent protein in the mammal, thereby alleviating the disease.

The invention further includes a method of alleviating a vitamin K-dependent protein associated disease in a mammal. The method comprises administering a vector comprising the chimeric nucleic acid of the invention to a mammal having the disease, wherein the chimeric nucleic acid is expressed in a cell in the mammal to produce a fully gamma carboxylated vitamin K-dependent protein in the mammal, thereby alleviating the disease.

Additionally, there is included a method of alleviating a vitamin K-dependent protein associated disease in a mammal. The method comprises administering a cell comprising the chimeric nucleic acid of the invention to a mammal having the disease, wherein the chimeric nucleic acid is expressed in the cell in the mammal to produce a fully gamma carboxylated vitamin K-dependent protein in the mammal, thereby alleviating the disease.

Further included is a pharmaceutical composition comprising the isolated chimeric nucleic acid of the invention, a pharmaceutical composition comprising a vector comprising the chimeric nucleic acid of the invention, a pharmaceutical composition comprising a cell comprising the chimeric nucleic acid of the invention, a pharmaceutical composition comprising the chimeric protein of the invention and a pharmaceutical composition comprising a fully carboxylated vitamin K-dependent protein.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1, comprising Figures 1Ai, 1Aii, 1Bi, 1Bii, and 1C, is a schematic representation of the cloning and expression of a fully gamma carboxylated Factor X protein. Figures 1Ai and 1Bi depict Round 1 PCR, and in both of these figures, "SSB" refers to signal sequence and propeptide. Figures 1Aii and 1Bii list the nucleotide sequences of prothrombin primers a and b (Figure 1Aii) and Factor X primers c and d (Figures 1Bii). In Figure 1Aii, { } indicates a BglII restriction site, [ ] indicates prothrombin sequence -43 to -39, ( ) indicates factor X

sequence 5 to 1, and < > indicates prothrombin sequence -1 to -5. In Figure 1Bii, the primer designated c is factor X sequence 1 to 8. Figure 1C is a flow sheet that depicts Round 2 PCR.

Figure 2 is a schematic representation of a protocol for purifying recombinant Factor X protein.

Figure 3 is a representation of the sequence alignment of the vitamin K-dependent coagulation factor propeptides. The eighteen amino acid propeptide sequence of the vitamin K-dependent coagulation factors, which is the primary binding site for the carboxylase, are shown above along with their inhibition constants toward a FIX propeptide/γ-carboxyglutamic acid substrate. The K_i values are a relative measure of the affinity of the propeptide for the carboxylase. These data are taken from Stanley et al., 1999, J. Biol. Chem. 274:16940. The sequences shown have the following SEQ ID NOS: Factor X (SEQ ID NO:1); Factor VII (SEQ ID NO:2); Protein S (SEQ ID NO:3); Factor IX (SEQ ID NO:4); Protein C (SEQ ID NO:5); Prothrombin (SEQ ID NO:6).

Figure 4, comprising Figures 4A-4C, is a series of graphs depicting separation of γ-carboxylated and uncarboxylated rFactor X by hydroxyapatite (HA) chromatography. HA chromatography was used to separate uncarboxylated and fully γ-carboxylated rFactor X. Figure 4A: Approximately 80% of the rwtFactor X eluted in the first peak and represents uncarboxylated protein, and the remaining 20% eluted in the second peak and represents fully γ-carboxylated protein. Figure 4B (clone B5) and Figure 4C (clone A1): rFactor X expressed with the prothrombin pre-pro-sequence also separated on HA into two peaks, with approximately 10% of the protein eluting in the peak 1 (uncarboxylated protein) and the remaining 90% of the protein eluting in the peak 2 (γ-carboxylated protein). Elution of each protein was monitored by absorbance a 280 nm (left axis).

Figure 5 is the nucleotide sequence which encodes the following propeptides: Factor X, Factor VII, Protein S, Factor IX, Protein C and Prothrombin (SEQ ID NOS: 7-12, respectively).

Figure 6, comprising Figures 6A-6Fii, is the nucleotide sequence encoding the signal, propeptide and mature protein sequence of the following

proteins: Factor X (Figure 6A), Factor VII (Figures 6Bi and 6Bii), Protein S (Figures 6Ci-6Ciii), Factor IX (Figures 6Di and 6Dii), Protein C (Figures 6Ei and 6Eii), and Prothrombin (Figures 6Fi and 6Fii; SEQ ID NOS: 13-18, respectively).

DETAILED DESCRIPTION OF THE INVENTION

The invention relates to the discovery of a method of generating a recombinant gamma carboxylated vitamin K-dependent protein, exemplified herein by Factor X , which is about 90 to 95% carboxylated. Thus, the invention provides a vast improvement over prior art methods wherein gamma carboxylated Factor X is produced which is only about 20 to 40% carboxylated (Larson et al., 1998, Biochemistry 37:5029-5038). Within the context of the present invention, a system has been designed which is applicable to and enhances the gamma carboxylation of vitamin K-dependent proteins.

Gamma carboxylation of vitamin K-dependent proteins occurs via binding of the carboxylase enzyme to the propeptide portion of the protein. It has been discovered in the present invention that it is possible to vary the affinity of the propeptide for the carboxylase. When the affinity of the propeptide is varied, the extent of gamma carboxylation of the mature protein is also varied. Thus, modification of the affinity of a propeptide for the carboxylase in a particular expression system may enhance or reduce gamma carboxylation of the protein. In the example presented herein, the affinity of the Factor X propeptide for carboxylase in HEK 293 cells was decreased. As a result, carboxylation of Factor X was enhanced from about 30% to 85%. Thus, the first step in enhancing gamma-carboxylation of a protein is to modify the affinity of the propeptide for the carboxylase. Depending on the expression system and the known affinity for that vitamin K-dependent protein, it may be necessary to either increase or decrease the affinity of the propeptide for the carboxylase. Each expression system may have an "optimal" propeptide affinity for its "own" carboxylase. It is a simple matter, once armed with the present invention to assess the effect of modification of any given propeptide on gamma carboxylation of any given mature vitamin K-dependent protein in any particular expression system.

While the vitamin K-dependent propeptides share sequence homology, their relative affinities for the carboxylase vary over a 100-fold, with the propeptide of Factor X binding with the greatest affinity followed by FVII, protein S, FIX, protein C, and prothrombin (Figure 3). As described herein, studies indicate that specific amino acids within these propeptide sequences are responsible for their reduced affinity for the carboxylase (Example 3). In the case of protein C and prothrombin which bind the carboxylase with similar affinities, a single amino acid change suffice to substantially increase their affinity for the carboxylase (Figure 3 and Example 3).

Although the examples provided herein include the use of a prothrombin propeptide sequence linked to a mature Factor X sequence, and a modified Factor X sequence linked to a mature Factor X sequence, the invention should in no way be construed as being limited to these examples, but rather, should be construed to include any and all combinations of propeptide sequences and mature vitamin K-dependent sequences which are presently known or become known. The use of any modified propeptide sequence that enhances gamma carboxylation of a mature vitamin K-dependent in any given expression system is contemplated in the invention.

According to the present invention, a chimeric cDNA has been generated in which DNA encoding a selected signal sequence and a propeptide sequence is fused to DNA encoding mature Factor X. Thus, the chimeric cDNA comprises a signal and propeptide sequence fused to a Factor X sequence. The type of signal sequence used is not important to the gamma carboxylation of the mature protein. Rather, as already noted herein, gamma carboxylation of the protein is affected by the propeptide sequence present in the construct comprising the chimeric nucleic acid. Thus, for purposes of the discussion which follows, the nature of the signal sequence is deemed to be irrelevant; however, it is assumed in all of the discussions which follow that a signal sequence is present in the nucleic acid sequence comprising the chimeric DNA of the invention. The signal sequence may be that of the

native protein, or it may be that of the propeptide sequence used, or it may be an unrelated signal sequence.

The propeptide sequence which is normally fused to DNA encoding the mature vitamin K-dependent protein, exemplified herein by Factor X, is replaced in the chimeric DNA of the invention by a new propeptide sequence. The new propeptide sequence is one which is not normally associated with the mature protein. A propeptide sequence which is normally associated with any given mature vitamin K-dependent protein is referred to herein as a "wild type" propeptide sequence. A wild type propeptide sequence is distinguished from the propeptide sequences useful in the present invention in that, in the latter case, the sequences have been altered such that the affinity of the propeptide for gamma carboxylase is different from that of the wild type propeptide.

Propeptide sequences useful in the invention therefore include altered forms of wild type sequences, and further include for example, a prothrombin propeptide sequence linked to a mature Factor X sequence, i.e., it is possible to practice the present invention using mixed and matched propeptide/mature vitamin K-dependent sequences. As noted herein, the propeptide sequence in vitamin K-dependent proteins is the recognition element for the enzyme which directs gamma carboxylation of the protein. Vitamin K-dependent proteins are not fully functional unless they comprise a high percentage of gamma carboxylated moieties. Thus, it is important when generating recombinant versions of these proteins that mechanisms be put in place to ensure full gamma carboxylation of the same. The replacement of the native propeptide sequence in a vitamin K-dependent protein by the prothrombin sequence results in a mature protein which comprises many more gamma carboxylated groups than that generated when the native propeptide sequence is used.

The invention therefore constitutes an improved strategy for producing highly gamma carboxylated recombinant vitamin K-dependent proteins. While the production of gamma carboxylated Factor X is exemplified herein, the invention should not be construed to be limited to the use of the method for production of highly

gamma carboxylated Factor X alone. Rather, the invention should be construed to encompass all vitamin K-dependent proteins, including, without limitation, Factor IX, Factor VII and protein C.

The sequence alignment of several propeptide sequences is shown in Figure 3. Thus, propeptides which are useful in the present invention are those which have the sequences shown in Figure 3 wherein an 18 amino acid sequence of several useful propeptides is shown along with the relative affinities of these propeptides for gamma carboxylase.

The corresponding nucleotide sequences which encode the propeptide sequences shown in Figure 3 are shown in Figure 5, and the nucleotide sequences encoding the entire proteins are shown in Figure 6.

One preferred propeptide for use in the present invention is prothrombin propeptide. Another preferred propeptide for use in the present invention is that of Factor X wherein the primary amino acid sequence of mature Factor X has been altered such that the altered Factor X has a lower affinity for gamma carboxylase. Altered Factor X propeptide having a lower affinity for gamma carboxylase than unaltered Factor X is termed "low affinity Factor X propeptide" herein. The expression of low affinity Factor X propeptide linked to mature Factor X protein results in the production of a higher concentration of Factor X which is highly gamma carboxylated compared with that produced when unaltered Factor X is used. While not wishing to be bound by theory, it is understood that when Factor X propeptide is high affinity Factor X propeptide, it binds tightly to gamma carboxylase during the process of carboxylation. Tightly bound gamma carboxylase is then unavailable for carboxylation of other Factor X molecules. On the other hand, low affinity Factor X propeptide binds gamma carboxylase less tightly, resulting in a higher turnover rate of the enzyme and therefore increased carboxylation of Factor X molecules in the reaction mixture.

The invention therefore also includes a low affinity propeptide, exemplified by the low affinity Factor X propeptide as described and defined herein.

A "low affinity propeptide" is one which has an affinity for gamma carboxylase which is similar to the affinity of prothrombin propeptide as disclosed herein.

Referring again to Figure 3 and Examples 2 and 3 herein, a low affinity propeptide may be generated by modifying any one of amino acids –9 or -13 on either prothrombin or protein C. Preferred modifications include the substitution of an Arg or a His residue at position –9 and the substitution of a Pro or a Ser residue at position –13.

The invention also includes a chimeric protein comprising a propeptide sequence fused to a mature vitamin K-dependent protein.

By the term "chimeric protein" as used herein is meant a propeptide sequence fused to a mature vitamin K dependent protein, wherein the combination of the two peptide sequences is a non-natural combination. For example, a Factor X propeptide sequence that has been altered using recombinant DNA technology which is fused to wild type mature Factor X is a non-natural combination. Similarly, a prothrombin propeptide fused to mature Factor X is a non-natural combination.

Preferred chimeric proteins include a propeptide selected from the group consisting of unaltered or altered Factor X, Factor VII, Protein S, Factor IX, Protein C and prothrombin, and a mature protein selected from the group consisting of Factor X, Factor VII, Protein S, Factor IX, Protein C and prothrombin. When the propeptide is an altered propeptide, then the altered propeptide can be linked to any one of the mature proteins listed herein. When the propeptide is unaltered propeptide, then the unaltered propeptide can be linked only to a mature protein with which it is not normally associated.

An "altered" propeptide is one in which at least one naturally occurring amino acid has been substituted with an amino acid which does not normally occur at that position in the propeptide.

The invention further includes a chimeric isolated nucleic acid encoding the chimeric protein of the invention. It is a simple matter once armed with the present invention, to construct a chimeric DNA comprising a signal and propeptide sequence

fused to a protein which depends upon gamma carboxylation for maximum biological activity. The nucleotide sequences which encode preferred propeptide sequences are shown in Figure 5 and the nucleotide sequences which encode the vitamin K-dependent proteins useful in the invention are shown in Figure 6.

By the term "chimeric DNA" as used in the context of the present invention, is meant a DNA encoding a signal sequence and a propeptide sequence fused to a DNA encoding a mature vitamin K dependent protein, wherein the combination of the two DNAs is a non-natural combination. For example, a Factor X propeptide sequence that has been altered using recombinant DNA technology which is fused to DNA encoding wild type mature Factor X is a non-natural combination. Similarly, DNA encoding prothrombin propeptide fused to DNA encoding mature Factor X is a non-natural combination.

The invention should be construed to include any and all chimeric DNAs and proteins encoded thereby, which comprise a signal and propeptide sequence fused to DNA encoding a gamma carboxylation-requiring protein.

Preferred chimeric DNAs include DNA encoding a propeptide selected from the group consisting of unaltered or altered Factor X, Factor VII, Protein S, Factor IX, Protein C and prothrombin, and DNA encoding a mature protein selected from the group consisting of Factor X, Factor VII, Protein S, Factor IX, Protein C and prothrombin. When the propeptide DNA encodes an altered propeptide, then the altered propeptide DNA can be linked to DNA encoding any one of the mature proteins listed herein. When the propeptide DNA encodes an unaltered propeptide, then the unaltered propeptide DNA can be linked only to DNA encoding a mature protein with which it is not normally associated.

The invention further includes a cell comprising the isolated chimeric nucleic acid of the invention, and a cell comprising the chimeric protein of the invention. The cell may be any type of cell including a prokaryotic and a eukaryotic cell. When the cell is a eukaryotic cell, the cell is preferably one in which gamma carboxylase is expressed. Suitable cells are described herein in the examples section,

although the invention should not be construed to be limited solely to the used of the specific cells exemplified herein.

Expression of the chimeric nucleic acid encoding the desired propeptide sequence fused to the desired vitamin K-dependent kinase sequence is effected by operably linking a suitable promoter sequence to the chimeric nucleic acid in such a manner that the promoter drives expression of the chimeric nucleic acid within a desired cell in the mammal. The type of promoter to be used will depend upon the expression system used to generate the protein and the invention should therefore not be construed to be limited to the use of any one type of promoter. Constitutive promoters, inducible promoters and tissue specific promoters are all contemplated as being useful in the present invention.

The invention also includes a method of optimizing the gamma carboxylation of a mature vitamin K-dependent protein. The method comprises substituting the naturally occurring propeptide sequence for the mature protein with an altered version of that propeptide sequence or by substituting the naturally occurring propeptide sequence for that protein with a different vitamin K-dependent protein propeptide sequence. Expression of a chimeric nucleic acid encoding a substituted or altered propeptide linked to a mature vitamin K-dependent protein in a cell results in the production of a propeptide having an altered affinity for gamma carboxylase. Gamma carboxylation of the protein then occurs via the action of gamma carboxylase in the cell. The effect of the substitution/alteration of the propeptide sequence on gamma carboxylation of the mature protein is assessed in any of the gamma carboxylation assays described herein. In this way, optimum gamma carboxylation of the protein can be achieved by substituting/altering the propeptide sequence fused thereto. Upon a reading of the present disclosure, it is a simple matter to generate "mix and match" propeptide/mature vitamin K-dependent chimeric DNAs or proteins, and to assess the level of gamma carboxylation of the mature protein.

The invention further includes a method of generating a fully carboxylated vitamin K-dependent kinase. The method comprises generating the

chimeric isolated nucleic acid of the invention, expressing the nucleic acid in a cell to produce the protein encoded thereby, wherein upon carboxylation of the protein by gamma carboxylase, the protein is fully carboxylated.

The term "fully gamma carboxylated protein" is used herein to refer to a protein wherein at least about 80% of the amino acids which should be gamma carboxylated are carboxylated. Preferably, at least about 85%, more preferably, at least about 90%, more preferably at least about 95% and even more preferably, at least about 99% of the amino acids which should be gamma carboxylated are gamma carboxylated.

The invention is useful for the production of quantities of sufficiently gamma carboxylated proteins for the use of the same as therapeutic molecules directly, or for their use in the development of small molecules which may be useful as agonists or antagonists of the subject protein. Further, the chimeric DNA of the invention is useful for in vivo production of the desired protein for treatment of vitamin K-dependent diseases or disorders.

Again, while the use of Factor X is exemplified herein, the invention should not be construed as being limited solely to the use of this protein and should be construed to include all vitamin K-dependent proteins and their use in treatment of disease states associated with them. With respect to Factor X, large quantities of highly gamma carboxylated forms of this molecule may be generated which may be useful directly as a therapeutic molecule for treatment of congenital or acquired Factor X deficiency, or it may be used for the development of small molecules which function as either agonists or antagonists of the native protein.

A pharmaceutical composition comprising fully gamma carboxylated mature vitamin K-dependent protein produced by the methods described herein, is also included in the invention.

In addition, the invention also includes a pharmaceutical composition comprising DNA encoding the chimeric protein of the invention as described in more detail elsewhere herein.

Further included in the invention is a method of alleviating a vitamin K-dependent kinase associated disease in a mammal. The mammal is preferably a human. In general, the vitamin K-dependent associated disease is one in which there is a deficiency of a vitamin K-dependent protein, which deficiency results in the disease state. The method comprises administering to the mammal a fully carboxylated mature vitamin K-dependent protein, wherein the carboxylated protein has been made according to the methods described herein.

Administration of a fully carboxylated protein to a mammal may also be effected by administering a chimeric nucleic acid encoding a propeptide linked to DNA encoding the mature vitamin K dependent protein. Preferably, the chimeric nucleic acid has a promoter operably linked thereto, wherein the promoter drives expression of the chimeric protein is a desired cell. Expression of the nucleic acid in a desired tissue in the mammal results in production of the mature protein in the tissue, which protein is gamma carboxylated by carboxylase in the tissue, thereby effecting administration of the protein to the mammal.

When the protein is administered to the mammal in the form of a chimeric nucleic acid encoding the same, the nucleic acid may be administered as naked DNA. However, preferably the nucleic acid is delivered to the mammal in the form of a vector, as that term is defined herein. Suitable vectors include both viral and non-viral vectors, the use of which is now well known in the art, and is described, for example, in Verma et al., 1997, Nature 389:239. It is well within the skill of the routineer in the field of the generation of and delivery of nucleic acids to mammals to determine exactly which vector to use and in what formulation and dosage, depending on the type of disease state to be alleviated.

As already noted herein, expression of the chimeric nucleic acid encoding the desired propeptide sequence fused to the desired vitamin K-dependent kinase sequence is effected by operably linking a suitable promoter sequence to the chimeric nucleic acid in such a manner that the promoter drives expression of the chimeric nucleic acid within a desired cell in the mammal. Depending on the disease

state to be treated, the promoter may be a constitutive promoter, an inducible promoter, and/or a tissue specific promoter.

When the fully carboxylated protein made by the methods of the invention is delivered to the mammal directly, the protein is formulated in a pharmaceutically acceptable carrier suitable for administration of the protein to the desired tissue in the mammal. Suitable pharmaceutical acceptable carriers include without limitation, saline, salts solution or other formulations apparent to those skilled in such administration. The chimeric protein may be administered to a mammal in one of the traditional modes (e.g., orally, parenterally, transdermally or transmucosally), in a sustained release formulation using a biodegradable biocompatible polymer, or by on-site delivery using micelles, gels and liposomes, or rectally (e.g., by suppository or enema) or nasally (e.g., by nasal spray). The appropriate pharmaceutically acceptable carrier will be evident to those skilled in the art and will depend in large part upon the route of administration.

The invention further includes a kit comprising a fully carboxylated vitamin K-dependent protein made by the methods described herein, and an instructional material for use of the kit.

Further included is a kit comprising the chimeric DNA of the invention and an instructional material for use of the kit.

The kit comprises the composition of the invention and an instructional material which describes adventitially administering the composition to a cell or a tissue of a mammal. In another embodiment, this kit comprises a (preferably sterile) solvent suitable for dissolving or suspending the composition of the invention prior to administering the compound to the mammal.

As used herein, an "instructional material" includes a publication, a recording, a diagram, or any other medium of expression which can be used to communicate the usefulness of the peptide of the invention in the kit for effecting alleviation of the various diseases or disorders recited herein. Optionally, or alternately, the instructional material may describe one or more methods of

alleviation the diseases or disorders in a cell or a tissue of a mammal. The instructional material of the kit of the invention may, for example, be affixed to a container which contains the peptide of the invention or be shipped together with a container which contains the peptide. Alternatively, the instructional material may be shipped separately from the container with the intention that the instructional material and the compound be used cooperatively by the recipient.

Definitions

As used herein, each of he following terms has the meaning associated with it in this section.

The articles "a" and "an" are used herein to refer to one or to more than one (i.e. to at least one) of the grammatical object of the article. By way of example, "an element" means one element or more than one element.

"Plurality" means at least two.

A "disease" is a state of health of an animal wherein the animal cannot maintain homeostasis, and wherein if the disease is not ameliorated then the animal's health continues to deteriorate. In contrast, a "disorder" in an animal is a state of health in which the animal is able to maintain homeostasis, but in which the animal's state of health is less favorable than it would be in the absence of the disorder. Left untreated, a disorder does not necessarily cause a further decrease in the animal's state of health.

A disease in a mammal is "alleviated" if the severity of a symptom of the disease, the frequency with which such a symptom is experienced by the mammal, or both, are reduced.

As used herein, the term "fused" is used to refer to nucleotide sequences which are directly attached to one another, having no intervening nucleotides. By way of example, the pentanucleotide 5'-AAAAA-3' is fused to the trinucleotide 5'-TTT-3' when the two are connected thus: 5'-AAAAATTT-3' or 5'-TTTAAAAA-3', but not when the two are connected thus: 5'-AAAAACTTT-3'.

As used herein, amino acids are represented by the full name thereof, by the three letter code corresponding thereto, or by the one-letter code corresponding thereto, as indicated in the following table:

Full Name	Three-Letter Code	One-Letter Code
Aspartic Acid	Asp	D
Glutamic Acid	Glu	E
Lysine	Lys	K
Arginine	Arg	R
Histidine	His	H
Tyrosine	Tyr	Y
Cysteine	Cys	C
Asparagine	Asn	N
Glutamine	Gln	Q
Serine	Ser	S
Threonine	Thr	T
Glycine	Gly	G
Alanine	Ala	A
Valine	Val	V
Leucine	Leu	L
Isoleucine	Ile	I
Methionine	Met	M
Proline	Pro	P
Phenylalanine	Phe	F
Tryptophan	Trp	W

As used herein, an "agonist" is a composition of matter which, when administered to a mammal such as a human, enhances or extends a biological activity attributable to the level or presence of an endogenous compound in the mammal.

An "antagonist" is a composition of matter which when administered to a mammal such as a human, inhibits a biological activity attributable to the level or presence of an endogenous compound in the mammal.

"Encoding" refers to the inherent property of specific sequences of nucleotides in a polynucleotide, such as a gene, a cDNA, or an mRNA, to serve as templates for synthesis of other polymers and macromolecules in biological processes having either a defined sequence of nucleotides (i.e., rRNA, tRNA and mRNA) or a defined sequence of amino acids and the biological properties resulting therefrom. Thus, a gene encodes a protein if transcription and translation of mRNA corresponding to that gene produces the protein in a cell or other biological system. Both the coding strand, the nucleotide sequence of which is identical to the mRNA sequence and is usually provided in sequence listings, and the non-coding strand, used as the template for transcription of a gene or cDNA, can be referred to as encoding the protein or other product of that gene or cDNA.

Unless otherwise specified, a "nucleotide sequence encoding an amino acid sequence" includes all nucleotide sequences that are degenerate versions of each other and that encode the same amino acid sequence. Nucleotide sequences that encode proteins and RNA may include introns.

An "isolated nucleic acid" refers to a nucleic acid segment or fragment which has been separated from sequences which flank it in a naturally occurring state, *e.g.*, a DNA fragment which has been removed from the sequences which are normally adjacent to the fragment, *e.g.*, the sequences adjacent to the fragment in a genome in which it naturally occurs. The term also applies to nucleic acids which have been substantially purified from other components which naturally accompany the nucleic acid, *e.g.*, RNA or DNA or proteins, which naturally accompany it in the cell. The term therefore includes, for example, a recombinant DNA which is incorporated into a vector, into an autonomously replicating plasmid or virus, or into the genomic DNA of a prokaryote or eukaryote, or which exists as a separate molecule (e.g, as a cDNA or a genomic or cDNA fragment produced by PCR or restriction enzyme digestion)

independent of other sequences. It also includes a recombinant DNA which is part of a hybrid gene encoding additional polypeptide sequence.

In the context of the present invention, the following abbreviations for the commonly occurring nucleic acid bases are used. "A" refers to adenosine, "C" refers to cytidine, "G" refers to guanosine, "T" refers to thymidine, and "U" refers to uridine.

A "polynucleotide" means a single strand or parallel and anti-parallel strands of a nucleic acid. Thus, a polynucleotide may be either a single-stranded or a double-stranded nucleic acid.

The term "nucleic acid" typically refers to large polynucleotides.

The term "oligonucleotide" typically refers to short polynucleotides, generally no greater than about 50 nucleotides. It will be understood that when a nucleotide sequence is represented by a DNA sequence (i.e., A, T, G, C), this also includes an RNA sequence (i.e., A, U, G, C) in which "U" replaces "T."

Conventional notation is used herein to describe polynucleotide sequences: the left-hand end of a single-stranded polynucleotide sequence is the 5'-end; the left-hand direction of a double-stranded polynucleotide sequence is referred to as the 5'-direction.

The direction of 5' to 3' addition of nucleotides to nascent RNA transcripts is referred to as the transcription direction. The DNA strand having the same sequence as an mRNA is referred to as the "coding strand"; sequences on the DNA strand which are located 5' to a reference point on the DNA are referred to as "upstream sequences"; sequences on the DNA strand which are 3' to a reference point on the DNA are referred to as "downstream sequences."

A "portion" of a polynucleotide means at least at least about twenty sequential nucleotide residues of the polynucleotide. It is understood that a portion of a polynucleotide may include every nucleotide residue of the polynucleotide.

"Recombinant polynucleotide" refers to a polynucleotide having sequences that are not naturally joined together. An amplified or assembled

recombinant polynucleotide may be included in a suitable vector, and the vector can be used to transform a suitable host cell.

A recombinant polynucleotide may serve a non-coding function (e.g., promoter, origin of replication, ribosome-binding site, etc.) as well.

A "recombinant polypeptide" is one which is produced upon expression of a recombinant polynucleotide.

"Polypeptide" refers to a polymer composed of amino acid residues, related naturally occurring structural variants, and synthetic non-naturally occurring analogs thereof linked via peptide bonds, related naturally occurring structural variants, and synthetic non-naturally occurring analogs thereof. Synthetic polypeptides can be synthesized, for example, using an automated polypeptide synthesizer.

The term "protein" typically refers to large polypeptides.

The term "peptide" typically refers to short polypeptides.

Conventional notation is used herein to portray polypeptide sequences: the left-hand end of a polypeptide sequence is the amino-terminus; the right-hand end of a polypeptide sequence is the carboxyl-terminus.

As used herein, a "native" protein or DNA molecule is one which is naturally occurring in a cell. "Naturally-occurring" as applied to an object refers to the fact that the object can be found in nature. For example, a polypeptide or polynucleotide sequence that is present in an organism (including viruses) that can be isolated from a source in nature and which has not been intentionally modified by man is naturally-occurring.

As used herein, the term "promoter/regulatory sequence" means a nucleic acid sequence which is required for expression of a gene product operably linked to the promoter/regulator sequence. In some instances, this sequence may be the core promoter sequence and in other instances, this sequence may also include an enhancer sequence and other regulatory elements which are required for expression of the gene product. The promoter/regulatory sequence may, for example, be one which expresses the gene product in a tissue specific manner.

A "constitutive promoter is a promoter which drives expression of a gene to which it is operably linked, in a constant manner in a cell. By way of example, promoters which drive expression of cellular housekeeping genes are considered to be constitutive promoters.

An "inducible" promoter is a nucleotide sequence which, when operably linked with a polynucleotide which encodes or specifies a gene product, causes the gene product to be produced in a living cell substantially only when an inducer which corresponds to the promoter is present in the cell.

A "tissue-specific" promoter is a nucleotide sequence which, when operably linked with a polynucleotide which encodes or specifies a gene product, causes the gene product to be produced in a living cell substantially only if the cell is a cell of the tissue type corresponding to the promoter.

By describing two polynucleotides as "operably linked" is meant that a single-stranded or double-stranded nucleic acid moiety comprises the two polynucleotides arranged within the nucleic acid moiety in such a manner that at least one of the two polynucleotides is able to exert a physiological effect by which it is characterized upon the other. By way of example, a promoter operably linked to the coding region of a gene is able to promote transcription of the coding region.

By the term "signal sequence" is meant a polynucleotide sequence which encodes a peptide that directs the path a polypeptide takes within a cell, i.e., it directs the cellular processing of a polypeptide in a cell, including, but not limited to, eventual secretion of a polypeptide from a cell. A signal sequence is a sequence of amino acids which are typically, but not exclusively, found at the amino terminus of a polypeptide which targets the synthesis of the polypeptide to the endoplasmic reticulum. In some instances, the signal peptide is proteolytically removed from the polypeptide and is thus absent from the mature protein.

A "prophylactic" treatment is a treatment administered to a subject who does not exhibit signs of a disease or exhibits only early signs of the disease for the purpose of decreasing the risk of developing pathology associated with the disease.

A "therapeutic" treatment is a treatment administered to a subject who exhibits signs of pathology for the purpose of diminishing or eliminating those signs.

A "vector" is a composition of matter which comprises an isolated nucleic acid and which can be used to deliver the isolated nucleic acid to the interior of a cell. Numerous vectors are known in the art including, but not limited to, linear polynucleotides, polynucleotides associated with ionic or amphiphilic compounds, plasmids, and viruses. Thus, the term "vector" includes an autonomously replicating plasmid or a virus. The term should also be construed to include non-plasmid and non-viral compounds which facilitate transfer of nucleic acid into cells, such as, for example, polylysine compounds, liposomes, and the like. Examples of viral vectors include, but are not limited to, adenoviral vectors, adeno-associated virus vectors, retroviral vectors, and the like.

"Expression vector" refers to a vector comprising a recombinant polynucleotide comprising expression control sequences operatively linked to a nucleotide sequence to be expressed. An expression vector comprises sufficient cis-acting elements for expression; other elements for expression can be supplied by the host cell or in an in vitro expression system. Expression vectors include all those known in the art, such as cosmids, plasmids (e.g., naked or contained in liposomes) and viruses that incorporate the recombinant polynucleotide.

The invention is now described with reference to the following example. This example is provided for the purpose of illustration only and the invention should in no way be construed as being limited to this example but rather should be construed to encompass any and all variations which become evident as a result of the teaching provided herein.

Example 1

Construction of Chimeric DNA and Production of Protein

Chimeric cDNA comprising DNA encoding prothrombin propeptide fused to mature Factor X DNA was constructed as described in Figure 1. The generation of gamma carboxylated Factor X protein was accomplished by transfection

of human embryonic kidney 293 cells with the chimeric DNA as described in Figure 2. The procedures for generation the chimeric DNAs of the invention are well known in the art and are described, for example, in Sambrook et al. (1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, New York) and in Ausubel et al. (1994, Current Protocols in Molecular Biology, John Wiley & Sons, New York).

The precise experimental protocols used for the generation of fully gamma carboxylated Factor X protein are described in Example 2.

Example 2. Enhanced Gamma Carboxylation of Recombinant Factor X Using the Prothrombin Propeptide

The present experiments were performed in order to optimize the expression of fully γ-carboxylated rFactor X and to determine the significance of the differential binding affinities of propeptide sequences for the γ-carboxylase. To this end, a chimeric Factor X cDNA harboring the prothrombin signal sequence and propeptide was created and transfected into HEK 293 cells. Stable transfectants were selected and expanded, and recombinant Factor X was purified and analyzed for γ-carboxyglutamic acid content. The data indicate that expression of rFactor X using the prothrombin signal sequence and propeptide results in much higher yields of fully γ-carboxylated material and is thus superior to the native Factor X signal sequence and propeptide. In addition, our results indicate that the affinity of the γ-carboxylase for the propeptide region can greatly influence the extent of γ-carboxylation. These data not only greatly facilitate the large scale production of functional rFactor X/FXa for detailed structure/function studies, but the observations are also be directly applicable to the production of other biologically active vitamin K-dependent proteins especially in a cellular setting where γ-carboxylation of this group of proteins is compromised.

The Materials and Methods used in this Example are now described.

All restriction enzymes were obtained from New England Biolabs, Beverly, MA. Pfu DNA polymerase was obtained from Stratagene, La Jolla, CA. Human embryonic kidney (HEK) 293 cells were obtained from ATCC, Rockville, MD. Lipofectamine, G418 (Geneticin), penicillin-streptomycin, trypsin-EDTA, L-

glutamine and DMEM F-12 were obtained from GIBCO-BRL, Gaithersburg, MD. Hydroxyapatite Bio Gel HT was obtained from BioRad Laboratories, Hercules, CA. Q-Sepharose was obtained from Pharmacia Biotech, Uppsala, Sweden. polyclonal Factor X and Factor X-horseradish peroxidase antibodies for ELISA were obtained from Dako, Carpinteria, CA. The calcium-dependent monoclonal human Factor X antibody (MoAb, 4G3) was obtained from Dr. Harold James, University of Texas, Tyler, TX.

Construction of the expression vector. In order to exchange the signal sequence and propeptide of Factor X with that of prothrombin the following specific oligonucleotide primers were constructed: primer A: 5'-GATCCAGATCTCCACCATGGCGCACGTCCGA-3' (SEQ ID NO:19), where the underlined portion is a BglII restriction site and the last 15 bases corresponds to prothrombin gene sequence coding for residues -41 to -37; primer B: 5'-AAGAAAGGAATTGGCTCGCCGGACCCGCTG-3' (SEQ ID NO:20), where the first 15 bases corresponds to Factor X gene sequence coding for residues +5 to +1 and the last 15 bases corresponds to prothrombin gene sequence coding for residues -1 to -5; primer C: 5'-GCCAATTCCTTTCTTGAAGAGATG-3' (SEQ ID NO:21), where the 24 bases correspond to Factor X gene sequence coding for residues +1 to +8; primer D: 5'-GAAACCCTCGTTTTCCTCATT-3' (SEQ ID NO:22), where the 21 bases corresponds to Factor X gene sequence coding for residues +220 to +214. The human prothrombin cDNA was kindly provided by Dr. Siriam Krishnaswamy, The Joseph Stokes Research Institute, Philadelphia, PA. The prothrombin DNA sequence encoding the signal sequence and propeptide was recombined with the DNA sequence of Factor X starting at position +1 by the technique of splicing by overlap extension or "geneSOEing", where primers B and C are the SOEing primers and primers A and D are the outside primers (Horton et al., 1989, Gene 77:61). The resulting 819 bp fragment was digested with BglII and SacII, gel purified and subcloned into pCMV4wt-Factor X. The new chimeric vector (pCMV4-ss-pro-II-Factor X) was transformed into competent bacterial cells, single colonies were picked, and vector

DNA was isolated by established techniques. To confirm the presence of the prothrombin signal sequence and propeptide and to ensure no polymerase-induced errors, the entire chimeric prothrombin/Factor X insert was subjected to DNA sequencing.

Expression of rFactor X. HEK 293 cells were transfected with pCMV4-ss-pro-II-Factor X using Lipofectamine according to the manufacturer's instructions. Cotransfection with a plasmid containing the neomycin resistance gene and the PACE/furin gene (pcDNA3-PACE; pcDNA3 was obtained from Invitrogen and the PACE cDNA was obtained from Genetics Institute, Boston, MA.) was performed at a 1:10 molar ratio (pcDNA3-PACE/pCMV4-ss-pro-II-Factor X). Transfectants were selected with the neomycin analogue G418, and resistant colonies were screened for Factor X production by sandwich ELISA as described (Larson et al., 1998, Biochemistry 37:5029). Selected clones were expanded into NUNC cell factories (1264 or 6320 cm^2; Nalge Nunc Int., Naperville, IL.) and a total of 6-15 liters of conditioned media was collected over 14-21 days. The medium was filtered and made 10 mM benzamidine prior to storage at -20°C.

Purification of rFactor X. rFactor X was purified from conditioned media using a three-step chromatographic approach (Q-Sepharose, Factor X immunoaffinity, and hydroxyapatite chromatography) essentially as described (Larson et al., 1998, Biochemistry 37:5029). The fully γ-carboxylated rFactor X eluting from the hydroxyapatite column was precipitated with ammonium sulfate and the protein was stored at –20°C in 50% glycerol/water. The concentration of rFactor X was determined by absorbance at 280 nm (M_r = 59,000; $E_{280nm}1\%$ = 11.6) (Di Scipio et al., 1977, Biochemistry 16:698).

Characterization of rFactor X. Protein purity was assessed using NuPAGE 4-12% Bis-Tris gels (Novex, San Diego, CA) followed by staining with Coomassie Brilliant Blue R-250. γ-Carboxyglutamic acid analysis was carried out according to the modified method of Price (Price 1983, Methods Enzymol. 91:13) for alkaline hydrolysis and separation of amino acids was accomplished using a DC-4A

cation exchange column on a Waters LC-1 Plus HPLC as described by Przysiecki (Przysiecki et al., 1987, Proc Natl. Acad Sci USA 84:7856). Known amounts of L-γ-carboxyglutamic acid (250 pmole) and L-aspartic acid (500 pmole) were used as standards for peak areas as well as retention times. The Gla and Asp/Asn peaks areas of the base hydrolyzed plasma-derived and rFactor X samples were compared to the peak areas of the Gla and Asp standards; moles of Gla per mole of protein were calculated from these values. Amino-terminal sequence analysis of both the heavy and light chains of rFactor X was accomplished by transferring these fragments to PVDF membranes (Matsudaira et al., 1987, J. Biol. Chem. 262:10035) followed by automatic Edman degradation on an Applied Biosystems 475A protein sequencing system (Kalafatis et al., 1993, J. Biol. Chem. 268:27246).

The Results of the experiments presented in this Example are now described.

Preparation and Expression of a Prothrombin/Factor X Chimera. The following experiment was conducted to determine if γ-carboxylation of rFactor X could be enhanced by exchanging its propeptide with one that binds the γ-carboxylase with a reduced affinity. Thus, the propeptide of Factor X was exchanged with that of prothrombin. Using the technique of splicing by overlap extension, the signal sequence and propeptide of prothrombin was attached to the Factor X cDNA starting at position +1 following three separate PCR reactions. The final PCR product was digested with BglII and SacII and was ligated into the mammalian expression vector pCMV4-wtFactor X. The entire insert containing the signal sequence and propeptide of prothrombin was verified by dideoxy sequencing. It should be noted that the signal sequence of prothrombin was included simply to facilitate PCR and subcloning of the prothrombin propeptide; it should not influence in any way the extent of γ-carboxylation.

Expression, Purification, and Characterization of a Recombinant Prothrombin Propeptide Factor X Chimeric Protein. The chimeric expression vector, pCMV4-ss-pro-II-wtFactor X, was used to transfect HEK 293 cells. Several clones

which were positive for rFactor X by ELISA were selected and subsequently expanded to establish cell lines. Two clones harboring the prothrombin pre-pro-sequence, clone B5 and clone A1, as well as native rwtFactor X (clone D3; described previously in Larson et al., 1998, Biochemistry 37:5029) are described in detail here . Each of the chimeric prothrombin propeptide rFactor X clones directed high level expression (B5; 4.0 μg/10^6cells/24 hour; A1; 2.3 μg/10^6cells/24 hour) compared with native rwtFactor X (D3; 4.0 μg/10^6cells/24 hour) indicating that the prothrombin pre-pro-sequence did not alter the ability of this cell system to express rFactor X. These cell lines were expanded into cell factories and conditioned media was collected over 14-21 days.

Purification of fully γ-carboxylated rFactor X from conditioned media was accomplished using a three-step chromatographic approach as previously described (Larson et al., 1998, Biochemistry 37:5029). Following Q-Sepharose chromatography for initial capture, rFactor X was purified by immunoaffinity chromatography using a monoclonal antibody (MoAb, 4G3; Kim et al., 1994, Biotechnol. Lett. 16:549) that binds all rFactor X and does not discriminate between uncarboxylated and fully γ-carboxylated protein. In order to separate these two forms of rFactor X, phosphate elution from hydroxyapatite was employed. It is known that that at low phosphate concentrations (~150 mM; peak 1) uncarboxylated rFactor X elutes (0-0.5 mole of Gla/mole of rFactor X), and at high phosphate concentrations (~275 mM; peak 2) fully γ-carboxylated material elutes (10.5-11.0 mole of Gla/mole of rFactor X) (Larson et al., 1998, Biochemistry 37:5029). Thus, elution of rFactor X from hydroxyapatite not only provides a useful way for isolating fully γ-carboxylated protein, but also enables us to determine how much of the total rFactor X produced by a given clone is fully γ-carboxylated.

The detailed purification table for clone B5 (rFactor X with the prothrombin pre-pro-sequence is presented in Table 1. Similar results were obtained using clone A1.

TABLE 1
Purification of rwtFX clone B5

Purification Step	Volume (mL)	rFX (mg/mL)	Total FX (mg)	% Recovery
Conditioned Media[a]	9660	0.003	27.5	100
Q-Sepharose[a]	150	0.18	27.0	98.2
Immunoaffinity (4G3)[b]	46.0	0.49	22.5	82.0
Hydroxyapatite Peak-1[b]	30.0	0.09	2.7	9.8
Hydroxyapatite Peak-2[b]	52.0	0.34	17.7	64.3
Final rFX Uncarboxylated [b,c]	0.6	3.25	2.0	7.1
Final rFX Carboxylated [b,c]	2.0	8.62	17.2	62.7

[a] As determined by FX specific ELISA
[b] As determined by absorbance at 280nm
[c] Final purified material following ammonium sulfate precipitation.

Chromatograms of the hydroxyapatite eluates of the prothrombin propeptide/rFactor X chimeras clones B5 and A1 are shown in Figures 2B and 2C, respectively. These data establish that approximately 90% of the total rFactor X applied to the column eluted in peak 2, indicating that the majority of the starting material was fully γ-carboxylated. This is in marked contrast to that seen with native rwtFactor X (Figure 2A) where only approximately 20% of the rFactor X was fully γ-carboxylated. Several (n = 7) rFactor X clones have also been expressed and purified having the prothrombin pre-pro-sequence and the results indicate that on average, 85% of the protein is fully γ-carboxylated, compared with approximately 35% of the material from rFactor X expressed with its native pre-pro-sequence (Table 2). In addition, clone B5 (with prothrombin pre-pro sequence) has been expanded, expressed, and rFactor X from this clone purified on three separate occasions having similar levels of γ–carboxyglutamic acid (approximately 90%) content obtained each time.

Table 2
Characterization of Various rFX Clones

+FX propeptide			+Prothrombin propeptide		
FX Clone	Expression (g/10^6cells /24 hr	% of Total rFX Fully - carboxylated	FX Clone	Expression (g/10^6cells/24 hr	% of Total rFX Fully - carboxylated
rFX-1 D3	4.0	20	rFX-1 B5	4.0	90
rFX-2 B5	3.0	30	rFX-2 A1	2.3	91
rFX-3 D1	4.0	40	rFX-3C5	1.0	89
rFX-4 C4	2.4	30	rFX-4 E2	2.3	82
rFX-5 A3	1.5	25	rFX-5 A4	1.0	81
rFX-6 B2	3.2	35	rFX-6 A6	1.2	73
rFX-7 C3	2.1	45	rFX-7 C1	0.5	91

Direct γ-carboxyglutamic acid analyses of the alkaline hydrolysate of rFactor X eluting from peaks 1 and 2 for both clones (B5 and A1) harboring the prothrombin pre-pro-sequence are presented in Table 3. Consistent with the previous findings, rFactor X eluting in the peak 1 was essentially uncarboxylated and material eluting in peak 2 was fully γ-carboxylated.

Table 3
-Carboxyglutamic Acid Analysis

Samples	Average ± SD mole Gla/mole protein	Theoretical
PD-h Prothrombin[a]	10.1± 0.6	10.0
PD-hFactor IX[a]	12.3±0.3	12.0
PD-hFactor X[a]	10.8±0.1	11.0
PD-h Thrombin [a,b]	ND	0
rwtFX-ss-pro-II (B5) Peak 1	0.1±0.02	0
rwtFX-ss-pro-II (B5) Peak 2	10.7±0.1	11.0
rwtFX-ss-pro-II (A1) Peak 1	0.2±0.04	0
rwtFX-ss-pro-II (A1) Peak 2	10.3±0.1	11.0

[a] Plasma-derived human coagulation factors used as standards. See Methods for determination
of Gla values. Values are the average of three separate determinations ± S.D.

[b] ND; No Gla peak was detectable.

Amino terminal sequence analysis of clone B5 (similar results obtained with clone A1) demonstrate that the prothrombin pre-pro-sequence was correctly processed from the rFactor X light chain irrespective of its γ-carboxyglutamic acid content (Table 4). In addition, blanks were obtained at positions 6, 7, 14, and 16 for rFactor X eluting in the second peak indicating the presence of γ-carboxyglutamic acid, whereas glutamic acid was present at these sites for rFactor X eluting in the first peak. These results are consistent with the γ-carboxyglutamic acid analyses. The rFactor X heavy chain from both peaks was also sequenced and the integrity of the heavy chain was confirmed.

Table 4

Amino-Terminal Sequence Analysis of rwtFX Clone B5					
Peak 1 From Hydroxyapatite Column			Peak 2 From Hydroxyapatite Column		
Cycle	Amino Acid	pmol	Cycle	Amino Acid	pmol
1	Ala	85.8	1	Ala	75.2
2	Asn	36.5	2	Asn	38.0
3	Ser	25.7	3	Ser	21.8
4	Phe	69.0	4	Phe	52.6
5	Leu	69.4	5	Leu	43.4
6	Glu	30.3	6	(Gla)	
7	Glu	51.4	7	(Gla)	
8	Met	54.6	8	Met	34.3
9	Lys	51.8	9	Lys	15.4
10	Lys	86.1	10	Lys	50.9
11	Gly	52.3	11	Gly	24.8
12	His	5.8	12	His	9.90
13	Leu	34.9	13	Leu	17.8
14	Glu	4.8	14	(Gla)	
15	Arg	25.0	15	Arg	13.8
16	Glu	12.2	16	(Gla)	
17	Cys	10.1	17	Cys	10.1
18	Met	10.2	18	Met	6.70

Following purification of fully γ-carboxylated rFactor X from hydroxyapatite, each of the rFactor X molecules as well as plasma-derived Factor X were subjected to SDS-PAGE. Each protein was judged to be >95% pure and migrated at the expected molecular weight under both reducing and non-reducing conditions. The amount of single chain rFactor X was <5% indicating almost complete removal of the internal tripeptide. Following activation of each recombinant protein by RVV-X and purification on benzamidine-sepharose, these proteins were determined to be identical to plasma-derived FXa with respect to clotting activity, chromogenic substrate activity, inhibitor sensitivity, prothrombinase activity, and cofactor binding.

Example 3. Amino acids Responsible for Reduced Affinities of Vitamin-K Dependent Propeptides for Gamma Carboxylase

As noted elsewhere herein, the binding of the gamma-glutamyl carboxylase to its protein substrates is mediated by a conserved 18 amino acid propeptide sequence found in all vitamin K-dependent proteins. It has recently been shown that the apparent affinities of the naturally occurring propeptides for the carboxylase vary over a 100-fold range and that the propeptide of bone Gla protein has severely impaired affinity for the carboxylase (Stanley et al., 1999 J. Biol. Chem. 274:16940-16944). In the present example, a consensus propeptide sequence that binds tighter (Ki = 0.43 nM) to the carboxylase than any known propeptide sequence is reported. Comparing the factor IX propeptide to the propeptides of protein C, bone Gla protein, and prothrombin, the weakest binding propeptides, facilitated the prediction of the residues responsible for the relative weak binding of these substrates to the carboxylase. Propeptides were then generated with the predicted amino acid changes and binding affinities of these propeptides was determined. The reduced binding affinity of these propeptides relative to that of FIX is due to residues -15 in protein C, -10 and -6 in bone Gla protein, and -9 in prothrombin. A role for the -9 position was not previously recognized but is further evidenced by the identification of a new, naturally occurring mutation at this position in factor IX which causes a warfarin-sensitive hemophilia B phenotype. In addition, it was discovered that

propeptides having mutations found in warfarin-sensitive patients have reduced affinity for the carboxylase, suggesting a physiological relevance of propeptide binding affinity.

The disclosures of each and every patent, patent application, and publication cited herein are hereby incorporated herein by reference in their entirety.

While this invention has been disclosed with reference to specific embodiments, it is apparent that other embodiments and variations of this invention may be devised by others skilled in the art without departing from the true spirit and scope of the invention. The appended claims are intended to be construed to include all such embodiments and equivalent variations.

What is claimed is:

1. An isolated chimeric nucleic acid comprising a nucleic acid sequence encoding a propeptide fused to a nucleic acid sequence encoding a vitamin K-dependent protein.

2. The isolated chimeric nucleic acid of claim 1, wherein said vitamin K-dependent protein is selected from the group consisting of Factor X, Factor VII, protein S, Factor IX, protein C and prothrombin.

3. The isolated chimeric nucleic acid of claim 1, wherein said propeptide is selected from the group consisting of altered or unaltered Factor X, Factor VII, protein S, Factor IX, protein C and prothrombin propeptide.

4. A vector comprising the isolated chimeric nucleic acid of claim 1.

5. A cell comprising the isolated chimeric nucleic acid of claim 1.

6. A chimeric protein comprising a propeptide fused to a vitamin K-dependent protein.

7. The chimeric protein of claim 6, wherein said vitamin K-dependent protein is selected from the group consisting of Factor X, Factor VII, protein S, Factor IX, protein C and prothrombin.

8. The chimeric protein of claim 6, wherein said propeptide is selected from the group consisting of altered or unaltered Factor X, Factor VII, protein S, Factor IX, protein C and prothrombin propeptide.

9. A cell comprising the chimeric protein of claim 6.

10. A method of optimizing the gamma carboxylation of a vitamin K-dependent protein, said method comprising introducing into a cell the chimeric nucleic acid of claim 1, expressing said chimeric nucleic in said cell, and assessing the level of gamma carboxylation of said vitamin K-dependent protein expressed by said chimeric nucleic acid, wherein the effect of the propeptide sequence on said gamma carboxylation is measured, said method further comprising modifying said nucleic acid encoding said propeptide until optimal gamma carboxylation of said vitamin K-dependent protein is achieved.

11. The method of claim 10, wherein said modifying includes substituting nucleic acid encoding said propeptide sequence with an altered or different propeptide sequence.

12. A method of producing a fully gamma carboxylated vitamin K-dependent protein, said method comprising introducing into a cell an isolated chimeric nucleic acid comprising a nucleic acid encoding a propeptide fused to a nucleic acid sequence encoding a vitamin K-dependent protein and expressing said protein therefrom, thereby producing a fully gamma carboxylated vitamin K-dependent protein.

13. A fully gamma carboxylated vitamin K-dependent protein made by the method of claim 12.

14. A method of alleviating a vitamin K-dependent protein associated disease in a mammal, said method comprising administering the protein of claim 13 to a mammal having said disease thereby alleviating said disease.

15. A method of alleviating a vitamin K-dependent protein associated disease in a mammal, said method comprising administering the isolated chimeric nucleic acid of claim 1 to a mammal having said disease, wherein said chimeric nucleic acid is expressed in a cell in said mammal to produce a fully gamma carboxylated vitamin K-dependent protein in said mammal, thereby alleviating said disease.

16. A method of alleviating a vitamin K-dependent protein associated disease in a mammal, said method comprising administering the vector of claim 4 to a mammal having said disease, wherein said chimeric nucleic acid is expressed in a cell in said mammal to produce a fully gamma carboxylated vitamin K-dependent protein in said mammal, thereby alleviating said disease.

17. A method of alleviating a vitamin K-dependent protein associated disease in a mammal, said method comprising administering the cell of claim 5 to a mammal having said disease, wherein said chimeric nucleic acid is expressed in said cell in said mammal to produce a fully gamma carboxylated vitamin K-dependent protein in said mammal, thereby alleviating said disease.

18. A pharmaceutical composition comprising the isolated chimeric nucleic acid of claim 1.

19. A pharmaceutical composition comprising the vector of claim 4.

20. A pharmaceutical composition comprising the cell of claim 5.

21. A pharmaceutical composition comprising the chimeric protein of claim 6.

22. A pharmaceutical composition comprising the protein of claim 13.

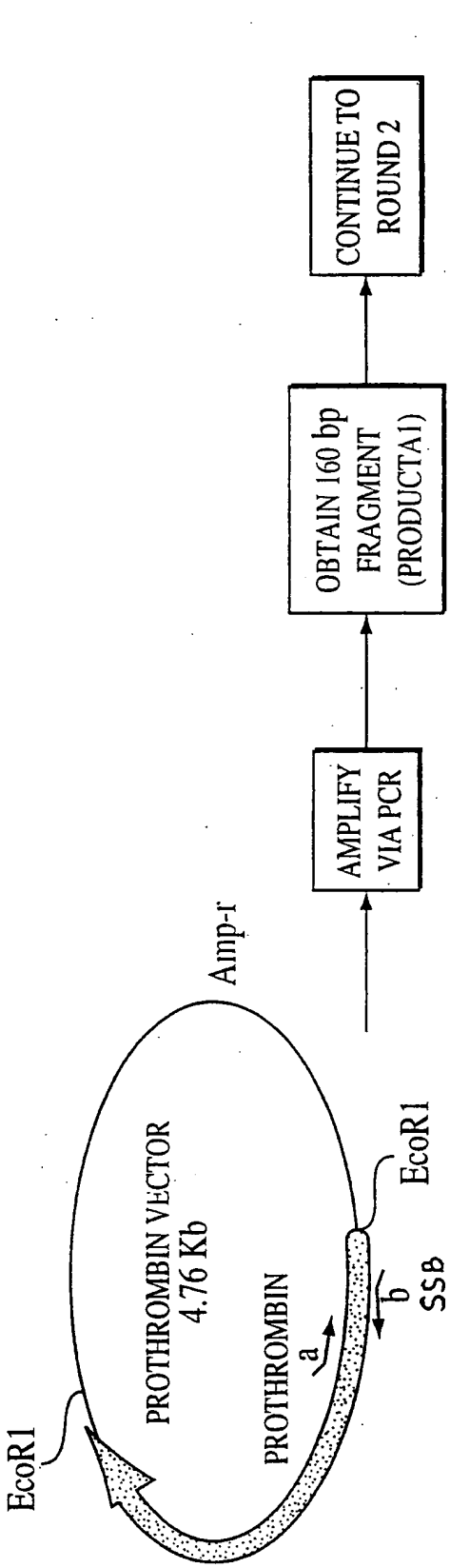

Fig. 1Ai

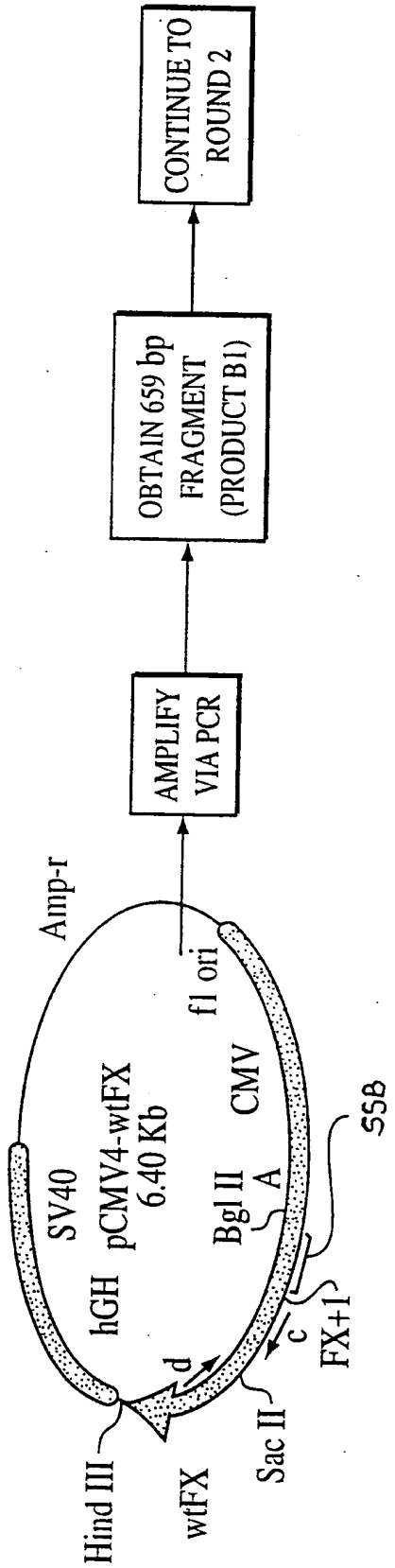

Fig. 1Bi

a. 5'-GAT CC{A GAT CT}C CAC C[AT GGC GCA CGT CCG A]-3'

b. 5'-(AAG AAA GGA ATT GGC) <TCG CCG GAC CCG CTG>-3'

FIG. 1Aii

c. 5'-GCC AAT TCC TTT CTT GAA GAG ATG-3'

d. 5'-GAA ACC CTC GTT TTC CTC ATT-3'

FIG. 1Bii

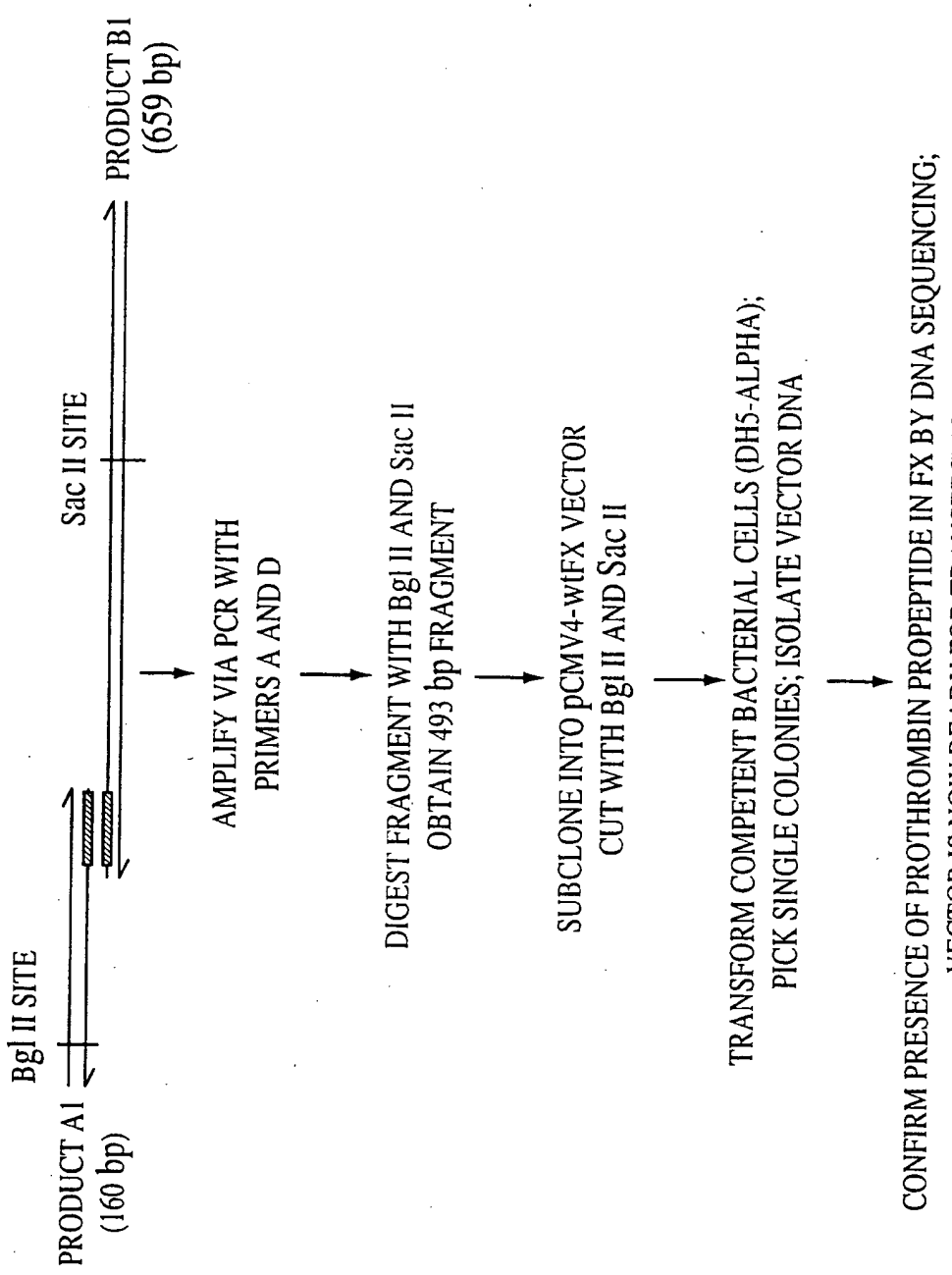
Bgl II SITE
PRODUCT A1
(160 bp)
Sac II SITE
PRODUCT B1
(659 bp)
AMPLIFY VIA PCR WITH
PRIMERS A AND D
DIGEST FRAGMENT WITH Bgl II AND Sac II
OBTAIN 493 bp FRAGMENT
SUBCLONE INTO pCMV4-wtFX VECTOR
CUT WITH Bgl II AND Sac II
TRANSFORM COMPETENT BACTERIAL CELLS (DH5-ALPHA);
PICK SINGLE COLONIES; ISOLATE VECTOR DNA
CONFIRM PRESENCE OF PROTHROMBIN PROPEPTIDE IN FX BY DNA SEQUENCING;
VECTOR IS NOW READY FOR TRANSFECTION

Fig. 1C

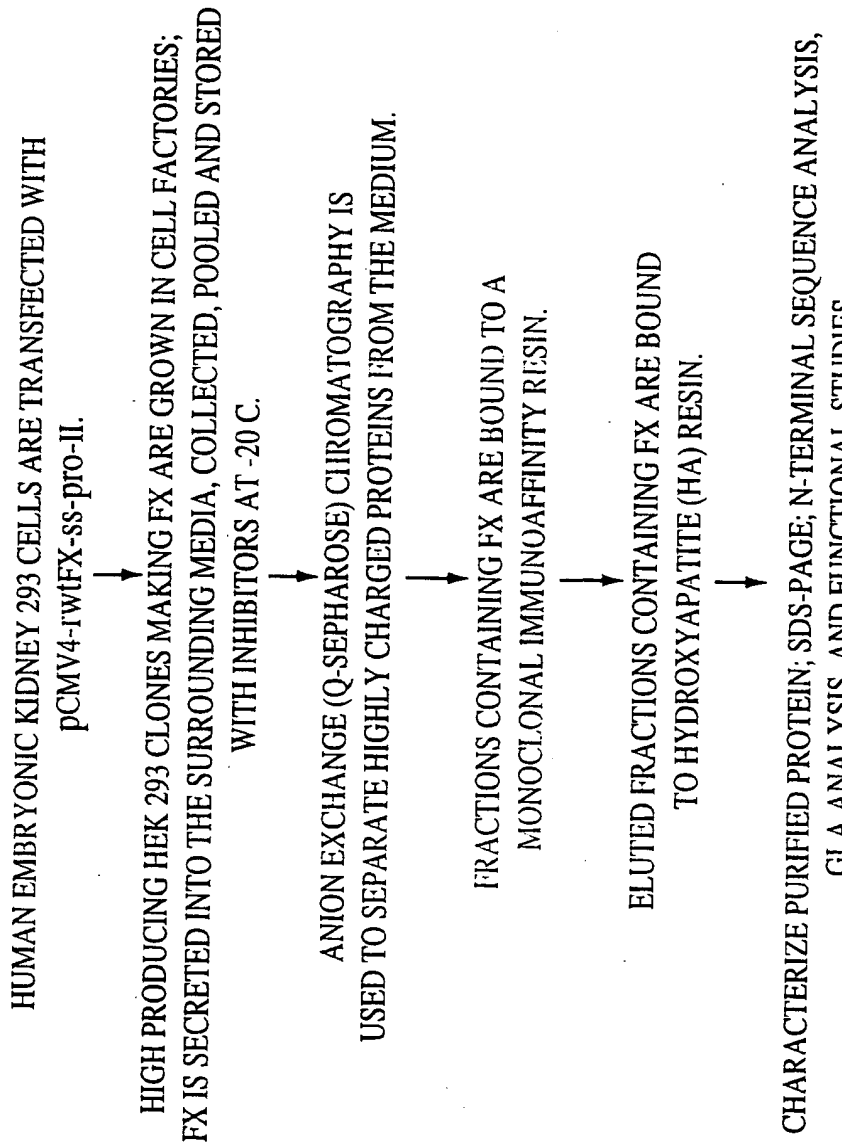
HUMAN EMBRYONIC KIDNEY 293 CELLS ARE TRANSFECTED WITH pCMV4-TwtFX-ss-pro-II.
HIGH PRODUCING HEK 293 CLONES MAKING FX ARE GROWN IN CELL FACTORIES; FX IS SECRETED INTO THE SURROUNDING MEDIA, COLLECTED, POOLED AND STORED WITH INHIBITORS AT -20 C.
ANION EXCHANGE (Q-SEPHAROSE) CHROMATOGRAPHY IS USED TO SEPARATE HIGHLY CHARGED PROTEINS FROM THE MEDIUM.
FRACTIONS CONTAINING FX ARE BOUND TO A MONOCLONAL IMMUNOAFFINITY RESIN.
ELUTED FRACTIONS CONTAINING FX ARE BOUND TO HYDROXYAPATITE (HA) RESIN.
CHARACTERIZE PURIFIED PROTEIN; SDS-PAGE; N-TERMINAL SEQUENCE ANALYSIS, GLA ANALYSIS, AND FUNCTIONAL STUDIES.

Fig. 2

	-18 -16 -10 -6 -1	Ki (nM)
FACTOR X	S L F I R R E Q A N N I L A R V T R	2.6
FACTOR VII	R V F V T E E E A H G V L H R R R R	11.1
PROTEIN S	A N F L S K Q Q A S Q V L V R K R R	12.2
FACTOR IX	R V F L D H E N A L K I L N R P K R	33.6
PROTEIN C	S V F S S S E R A H Q V L R I R K R	230
PROTHROMBIN	H V F L A P Q Q A R S L L Q R V R R	277

Fig. 3

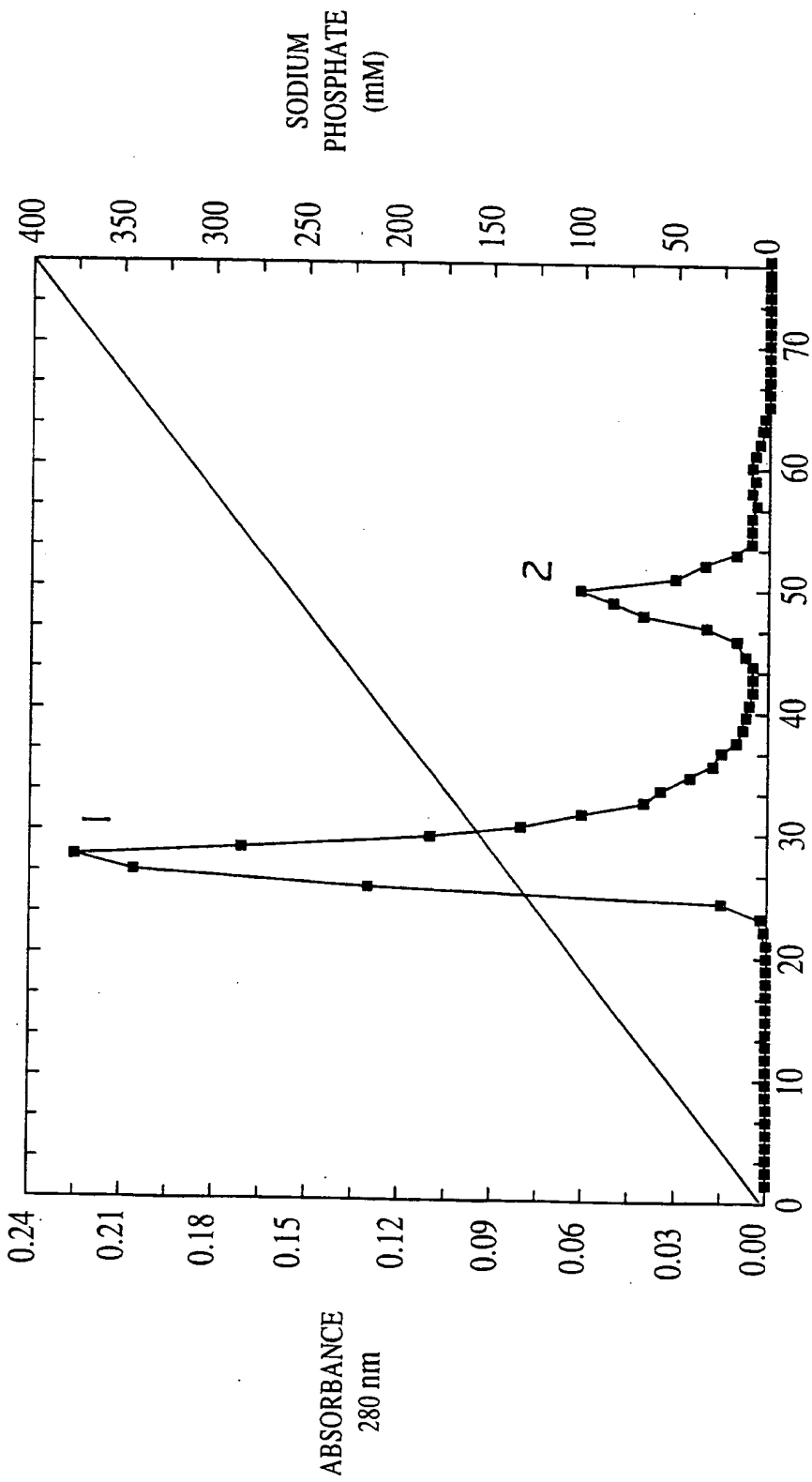
SODIUM PHOSPHATE (mM)
400
350
300
250
200
150
100
50
0
ABSORBANCE 280 nm
0.24
0.21
0.18
0.15
0.12
0.09
0.06
0.03
0.00
FRACTION #
0
10
20
30
40
50
60
70
1
2

Fig. 4A

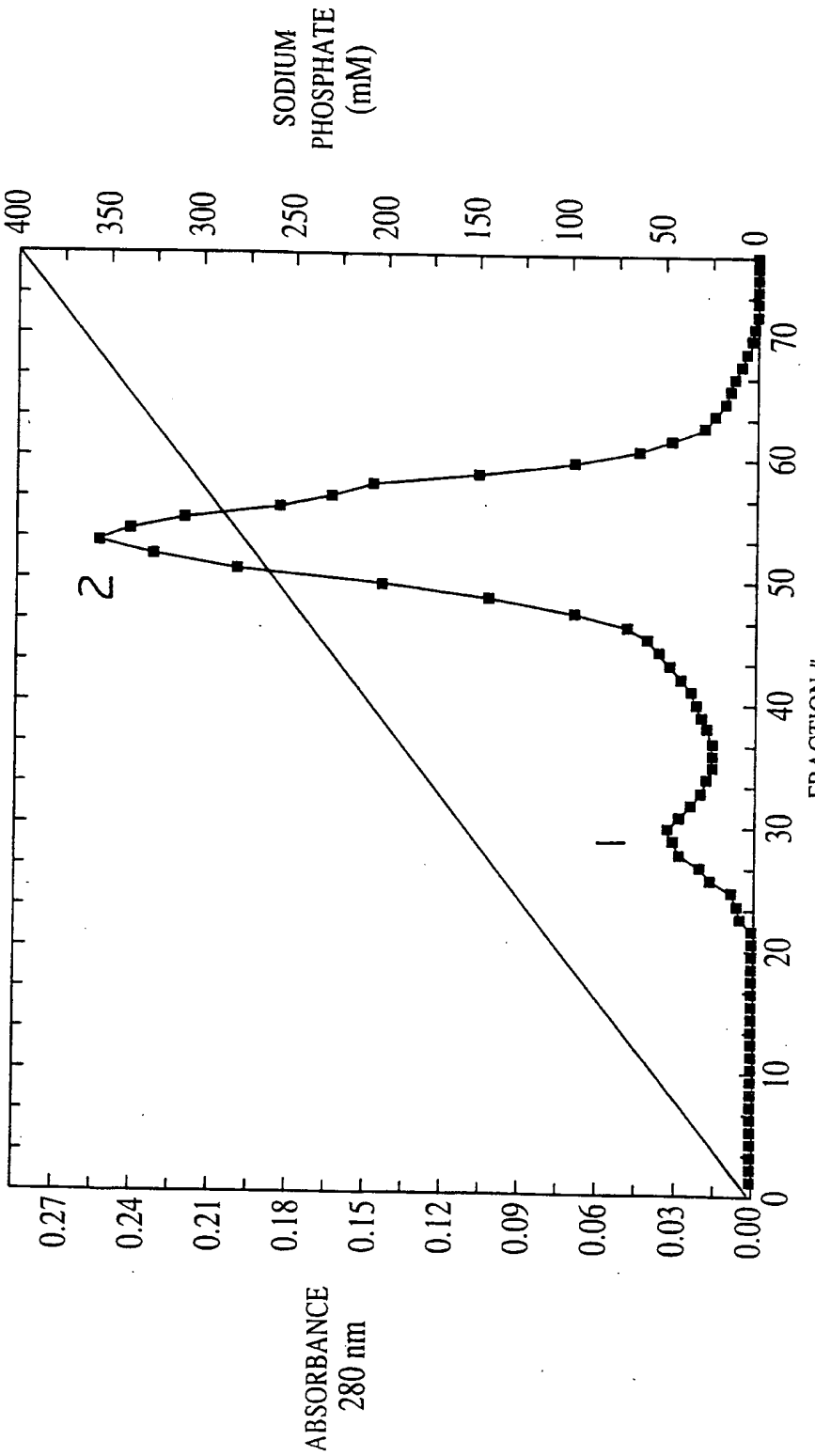
SODIUM PHOSPHATE (mM)
400
350
300
250
200
150
100
50
0
ABSORBANCE 280 nm
0.27
0.24
0.21
0.18
0.15
0.12
0.09
0.06
0.03
0.00
0
10
20
30
40
50
60
70
FRACTION #
1
2

Fig. 4B

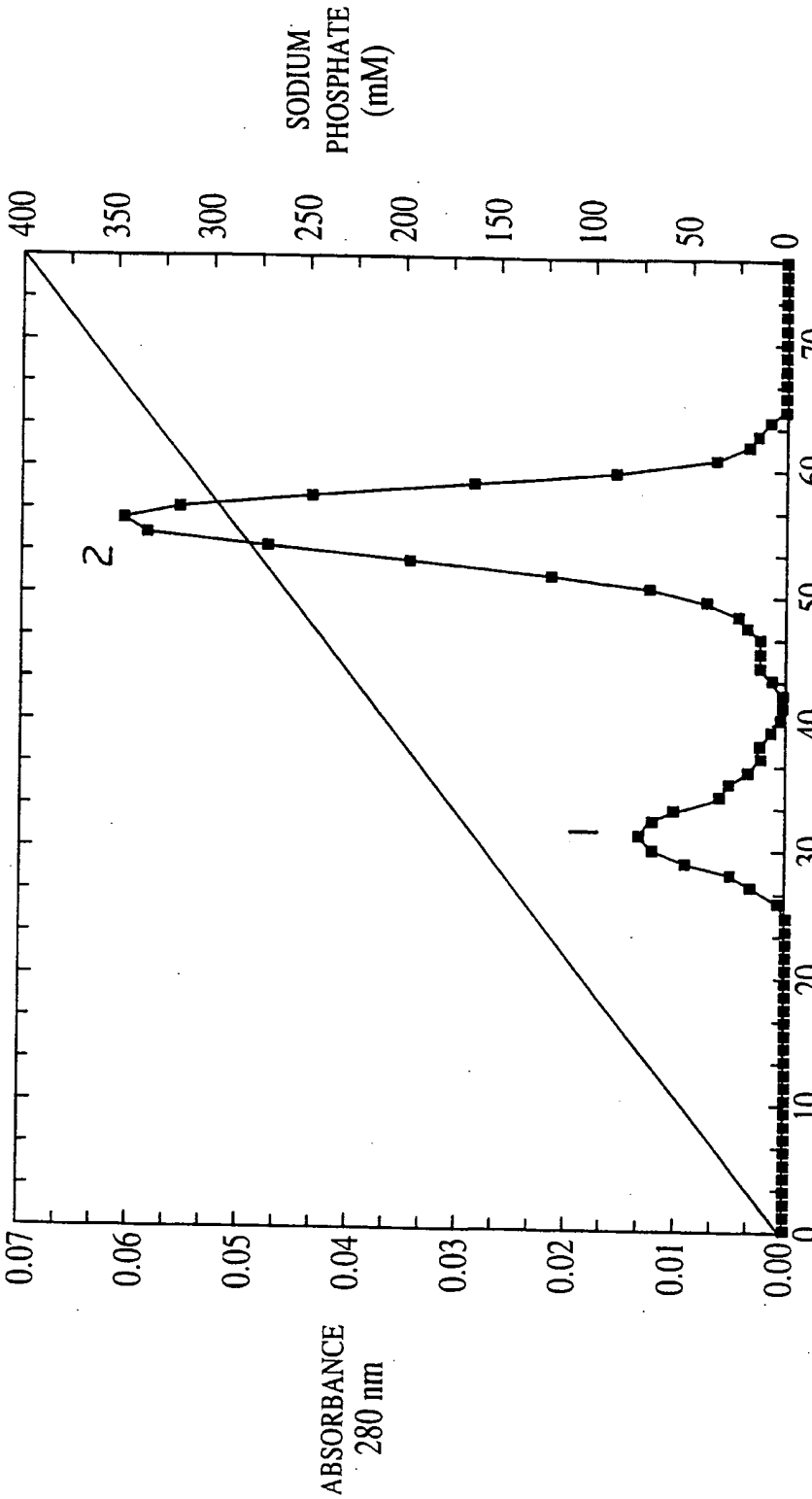
SODIUM PHOSPHATE (mM)
400
350
300
250
200
150
100
50
0
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00
ABSORBANCE 280 nm
0
10
20
30
40
50
60
70
FRACTION #
1
2

Fig. 4C

FIG. 5

Factor X (66-120):	AGTCTGTTCATCCGCAGGGAGCAGGCCAACAACATCCTGGCGAGGGTCACGAGG
Factor VII (161-215):	AGAGTCTTCGTAACCCAGGAGGAAGCCCACGGCGTCCTGCACCGGCGCCGGCGC
Protein S (216-269):	GCAAACCTTCTGTCAAAGCAACAGGCTTCACAAGTCCTGGTTAGGAAGCGTCGT
Factor IX (85-138):	ACAGTTTTTCTTGATCATGAAAACGCCAACAAAATTCTGAATCGGCCAAAGAGG
Protein C (170-223):	TCAGTGTTCTCCAGCAGCGAGCGTGCCCACCAGGTGCTGCGGATCCGCAAACGT
Prothrombin (107-160):	CATGTGTTCCTGGCTCCTCAGCAAGCACGGTCGCTGCTCCAGCGGGTCCGGCGA

FIG. 6A

```
   1 ATGGGGCGCC CACTGCACCT CGTCCTGCTC AGTGCCTCCC TGGCTGGCCT CCTGCTGCTC
  61 GGGGAAAGTC TGTTCATCCG CAGGGAGCAG GCCAACAACA TCCTGGCGAG GGTCACGAGG
 121 GCCAATTCCT TTCTTGAAGA GATGAAGAAA GGACACCTCG AAAGAGAGTG CATGGAAGAG
 181 ACCTGCTCAT ACGAAGAGGC CCGCGAGGTC TTTGAGGACA GCGACAAGAC GAATGAATTC
 241 TGGAATAAAT ACAAAGATGG CGACCAGTGT GAGACCAGTC CTTGCCAGAA CCAGGGCAAA
 301 TGTAAAGACG GCCTCGGGGA ATACACCTGC ACCTGTTTAG AAGGATTCGA AGGCAAAAAC
 361 TGTGAATTAT TCACACGGAA GCTCTGCAGC CTGGACAACG GGGACTGTGA CCAGTTCTGC
 421 CACGAGGAAC AGAACTCTGT GGTGTGCTCC TGCGCCCGCG GGTACACCCT GGCTGACAAC
 481 GGCAAGGCCT GCATTCCCAC AGGGCCCTAC CCCTGTGGGA AACAGACCCT GGAACGCAGG
 541 AAGAGGTCAG TGGCCCAGGC CACCAGCAGC AGCGGGGAGG CCCCTGACAG CATCACATGG
 601 AAGCCATATG ATGCAGCCGA CCTGGACCCC ACCGAGAACC CCTTCGACCT GCTTGACTTC
 661 AACCAGACGC AGCCTGAGAG GGGCGACAAC AACCTCACCA GGATCGTGGG AGGCCAGGAA
 721 TGCAAGGACG GGGAGTGTCC CTGGCAGGCC CTGCTCATCA ATGAGGAAAA CGAGGGTTTC
 781 TGTGGTGGAA CTATTCTGAG CGAGTTCTAC ATCCTAACGG CAGCCCACTG TCTCTACCAA
 841 GCCAAGAGAT TCAAGGTGAG GGTAGGGGAC CGGAACACGG AGCAGGAGGA GGGCGGTGAG
 901 GCGGTGCACG AGGTGGAGGT GGTCATCAAG CACAACCGGT TCACAAAGGA GACCTATGAC
 961 TTCGACATCG CCGTGCTCCG GCTCAAGACC CCCATCACCT TCCGCATGAA CGTGGCGCCT
1021 GCCTGCCTCC CCGAGCGTGA CTGGGCCGAG TCCACGCTGA TGACGCAGAA GACGGGGATT
1081 GTGAGCGGCT TCGGGCGCAC CCACGAGAAG GGCCGGCAGT CCACCAGGCT CAAGATGCTG
1141 GAGGTGCCCT ACGTGGACCG CAACAGCTGC AAGCTGTCCA GCAGCTTCAT CATCACCCAG
1201 AACATGTTCT GTGCCGGCTA CGACACCAAG CAGGAGGATG CCTGCCAGGG GGACAGCGGG
1261 GGCCCGCACG TCACCCGCTT CAAGGACACC TACTTCGTGA CAGGCATCGT CAGCTGGGGA
1321 GAGGGCTGTG CCCGTAAGGG GAAGTACGGG ATCTACACCA AGGTCACCGC CTTCCTCAAG
1381 TGGATCGACA GGTCCATGAA AACCAGGGGC TTGCCCCAAGG CCAAGAGCCA TGCCCCGGAG
1441 GTCATAACGT CCTCTCCATT AAAGTGAGAT CCCACTCAAA AAAAAAAAAA AAAAAAAAAA
1501 AAAAAA
```

FIG. 6Bi

```
   1 TCAACAGGCA GGGGCAGCAC TGCAGAGATT TCATCATGGT CTCCCAGGCC CTCAGGCTCC
  61 TCTGCCTTCT GCTTGGGCTT CAGGGCTGCC TGGCTGCAGG CGGGGTCGCT AAGGCCTCAG
 121 GAGGAGAAAC ACGGGACATG CCGTGGAAGC CGGGGCCTCA CAGAGTCTTC GTAACCCAGG
 181 AGGAAGCCCA CGGCGTCCTG CACCGGCGCC GGCGCGCCAA CGCGTTCCTG GAGGAGCTGC
 241 GGCCGGGCTC CCTGGAGAGG GAGTGCAAGG AGGAGCAGTG CTCCTTCGAG GAGGCCCGGG
 301 AGATCTTCAA GGACGCGGAG AGGACGAAGC TGTTCTGGAT TTCTTACAGT GATGGGACC
 361 AGTGTGCCTC AAGTCCATGC CAGAATGGGG GCTCCTGCAA GGACCAGCTC CAGTCCTATA
 421 TCTGCTTCTG CCTCCCTGCC TTCGAGGGCC GGAACTGTGA GACGCACAAG GATGACCAGC
 481 TGATCTGTGT GAACGAGAAC GGCGGCTGTG AGCAGTACTG CAGTGACCAC ACGGGCACCA
 541 AGCGCTCCTG TCGGTGCCAC GAGGGGTACT CTCTGCTGGC AGACGGGGTG TCCTGCACAC
 601 CCACAGTTGA ATATCCATGT GGAAAAATAC CTATTCTAGA AAAAGAAAT GCCAGCAAAC
 661 CCCAAGGCCG AATTGTGGGG GGCAAGGTGT GCCCCAAAGG GGAGTGTCCA TGGCAGGTCC
 721 TGTTGTTGGT GAATGGAGCT CAGTTGTGTG GGGGGACCCT GATCAACACC ATCTGGGTGG
 781 TCTCCGCGGC CCACTGTTTC GACAAAATCA AGAACTGGAG GAACCTGATC GCGGTGCTGG
 841 GCGAGCACGA CCTCAGCGAG CACGACGGGG ATGAGCAGAG CCGGCGGGTG GCGCAGGTCA
 901 TCATCCCCAG CACGTACGTC CCGGGCACCA CCAACCACGA CATCGCGCTG CTCCGCCTGC
 961 ACCAGCCCGT GGTCCTCACT GACCATGTGG TGCCCCTCTG CCTGCCCGAA CGGACGTTCT
1021 CTGAGAGGAC GCTGGCCTTC GTGCGCTTCT CATTGGTCAG CGGCTGGGGC CAGCTGCTGG
1081 ACCGTGGCGC CACGGCCCTG GAGCTCATGG TGCTCAACGT GCCCCGGCTG ATGACCCAGG
1141 ACTGCCTGCA GCAGTCACGG AAGGTGGGAG ACTCCCCAAA TATCACGGAG TACATGTTCT
1201 GTGCCGGCTA CTCGGATGGC AGCAAGGACT CCTGCAAGGG GGACAGTGGA GGCCCACATG
1261 CCACCCACTA CCGGGGCACG TGGTACCTGA CGGGCATCGT CAGCTGGGGC CAGGGCTGCG
1321 CAACCGTGGG CCACTTTGGG GTGTACACCA GGGTCTCCCA GTACATCGAG TGGCTGCAAA
1381 AGCTCATGCG CTCAGAGCCA CGCCCAGGAG TCCTCCTGCG AGCCCCATTT CCCTAGCCCA
1441 GCAGCCCTGG CCTGTGGAGA GAAAGCCAAG GCTGCGTCGA ACTGTCCTGG CACCAAATCC
1501 CATATATTCT TCTGCAGTTA ATGGGGTAGA GGAGGGCATG GGAGGGAGGG AGAGGTGGGG
1561 AGGGAGACAG AGACAGAAAC AGAGAGAGAC AGAGACAGAG AGAGACTGAG GGAGAGACTC
```

FIG. 6Bii

```
1621 TGAGGACATG GAGAGAGACT CAAAGAGACT CCAAGATTCA AAGAGACTAA TAGAGACACA
1681 GAGATGGAAT AGAAAAGATG AGAGGCAGAG GCAGACAGGC GCTGGACAGA GGGGCAGGGG
1741 AGTGCCAAGG TTGTCCTGGA GGCAGACAGC CCAGCTGAGC CTCCTTACCT CCCTTCAGCC
1801 AAGCCCCACC TGCACGTGAT CTGCTGGCCC TCAGGCTGCT GCTCTGCCTT CATTGCTGGA
1861 GACAGTAGAG GCATGAACAC ACATGGATGC ACACACACAC ACGCCCAATGC ACACACACAG
1921 AGATATGCAC ACACACGGAT GCACACACAG ATGGTCACAC AGAGATACGC AAACACACCG
1981 ATGCACACGC ACATAGAGAT ATGCACACAC AGATGCACAC ACAGATATAC ACATGGATGC
2041 ACGCACATGC CAATGCACGC ACACATCAGT GCACACGGAT GCACAGAGAT ATGCACACAC
2101 CGATGTGCGC ACACACAGAT ATGCACACAC ATGGATGAGC ACACACACAC CAAGTGCGCA
2161 CACACACCGA TGTACACACA CAGATGCACA CACAGATGCA CACACACCGA TGCTGACTCC
2221 ATGTGTGCTG TCCTCTGAAG GCGGTTGTTT AGCTCTCACT TTTCTGGTTC TTATCCATTA
2281 TCATCTTCAC TTCAGACAAT TCAGAAGCAT CACCATGCAT GGTGGCGAAT GCCCCCAAAC
2341 TCTCCCCCAA ATGTATTTCT CCCTTCGCTG GGTGCCGGGC TGCACAGACT ATTCCCCACC
2401 TGCTTCCCAG CTTCACAATA AACGGCTGCG TCTCCTCCGC ACACCTGTGG TGCCTGCCAC
2461 CC
```

FIG. 6Ci

```
   1 CTGCAGGGGG GGGGGGGGGG GGGGGGGGCG CAGCACGGCT CAGACCGAGG
  61 CGCACAGGCT CGCAGCTCCG GGCGCCTAGC GCCCGGTCCC CGCCGCGACG CGCCACCGTC
 121 CCTGCCGGCG CCTCCGCGCC TTCGAAATGA GGGTCCTGGG TGGGCGCTGC GGGGCGCCGC
 181 TGGCGTGTCT CCTCCTAGTG CTTCCCGTCT CAGAGGCAAA CCTTCTGTCA AAGCAACAGG
 241 CTTCACAAGT CCTGGTTAGG AAGCGTCGTG CAAATTCTTT ACTTGAAGAA ACCAAACAGG
 301 GTAATCTTGA AAGAGAATGC ATCGAAGAAC TGTGCAATAA AGAAGAAGCC AGGGAGGTCT
 361 TTGAAAATGA CCCGGAAACG GATTATTTTT ATCCAAAATA CTTAGTTTGT CTTCGCTCTT
 421 TTCAAACTGG GTTATTCACT GCTGCACGTC AGTCAACTAA TGCTTATCCT GACCTAAGAA
 481 GCTGTGTCAA TGCCATTCCA GACCAGTGTA GTCCTCTGCC ATGCAATGAA GATGGATATA
 541 TGAGCTGCAA AGATGGAAAA GCTTCTTTTA CTTGCACTTG TAAACCAGGT TGGCAAGGAG
```

FIG. 6Cii

```
601  AAAAGTGTGA ATTTGACATA AATGAATGCA AAGATCCCTC AAATATAAAT GGAGGTTGCA
661  GTCAAATTTG TGATAATACA CCTGGAAGTT ACCACTGTTC CTGTAAAAAT GGTTTTGTTA
721  TGCTTTCAAA TAAGAAAGAT TGTAAAGATG TGGATGAATG CTCTTTGAAG CCAAGCATTT
781  GTGGCACAGC TGTGTGCAAG AACATCCCAG GAGATTTTGA ATGTGAATGC CCCGAAGGCT
841  ACAGATATAA TCTCAAATCA AAGTCTTGTG AAGATATAGA TGAATGCTCT GAGAACATGT
901  GTGCTCAGCT TTGTGTCAAT TACCCTGGAG GTTACACTTG CTATTGTGAT GGGAAGAAAG
961  GATTCAAACT TGCCCAAGAT CAGAAGAGTT GTGAGGTTGT TTCAGTGTGC CTTCCCTTGA
1021 ACCTTGACAC AAAGTATGAA TTACTTTACT TGGCGGAGCA GTTTGCAGGG GTTGTTTTAT
1081 ATTTAAAATT TCGTTTGCCA GAAATCAGCA GATTTTCAGC AGAATTTGAT TTCCGGACAT
1141 ATGATTCAGA AGGCGTGATA CTGTACGCAG AATCTATCGA TCACTCAGCG TGGCTCCTGA
1201 TTGCACTTCG TGGTGGAAAG ATTGAAGTTC AGCTTAAGAA TGAACATACA TCCAAAATCA
1261 CAACTGGAGG TGATGTTATT AATAATGGTC TATGGAATAT GGTGTCTGTG GAAGAATTAG
1321 AACATAGTAT TAGCATTAAA ATAGCTAAAG AAGCTGTGAT GGATATAAAT AAACCTGGAC
1381 CCCTTTTTAA GCCGGAAAAT GGATTGCTGG AAACCAAAGT ATACTTTGCA GGATTCCCTC
1441 GGAAAGTGGA AAGTGAACTC ATTAAACCGA TTAACCCTCG TCTAGATGGA TGTATACGAA
1501 GCTGGAATTT GATGAAGCAA GGAGCTTCTG GAATAAAGGA AATTATTCAA GAAAAACAAA
1561 ATAAGCATTG CCTGGTTACT GTGGAGAAGG GCTCCTACTA TCCTGGTTCT GGAATTGCTC
1621 AATTTCACAT AGATTATAAT AATGTATCCA GTGCTGAGGG TTGGCATGTA AATGTGACCT
1681 TGAATATTCG TCCATCCACG GGCACTGGTG TTATGCTTGC CTTGGTTTCT GGTAACAACA
1741 CAGTGCCCTT TGCTGTGTCC TTGGTGGACT CCACCTCTGA AAAATCACAG GATATTCTGT
1801 TATCTGTTGA AAATACTGTA ATATATCGGA TACAGGCCCT AAGTCTATGT TCCGATCAAC
1861 AATCTCATCT GGAATTTAGA GTCAACAGAA ACAATCTGGA GTTGTCGACA CCACTTAAAA
1921 TAGAAACCAT CTCCCATGAA GACCTTCAAA GACAACTTGC CGTCTTGGAC AAAGCAATGA
1981 AAGCAAAAGT GGCCACATAC CTGGGTGGCC TTCCAGATGT TCCATTCAGT GCCACACCAG
2041 TGAATGCCTT TTATAATGGC TGCATGGAAG TGAATATTAA TGGTGTACAG TTGGATCTGG
2101 ATGAAGCCAT TTCTAAACAT AATGATATTA GAGCTCACTC ATGTCCATCA GTTTGGAAAA
2161 AGACAAAGAA TTCTTAAGGC ATCTTTTCTC TGCTTATAAT ACCTTTTCCT TGTGTGTAAT
```

FIG. 6Ciii

```
2221 TATACTTATG TTTCAATAAC AGCTGAAGGG TTTTATTTAC AATGTGCAGT CTTTGATTAT
2281 TTTGTGGTCC TTTCCTGGGA TTTTTAAAAG GTCCTTTGTC AAGGAAAAAA ATTCTGTTGT
2341 GATATAAATC ACAGTAAAGA AATTCTTACT TCTCTTGCTA TCTAAGAATA GTGAAAAATA
2401 ACAATTTTAA ATTTGAATTT TTTTCCTACA AATGACAGTT TCAATTTTTG TTTGTAAAAC
2461 TAAATTTTAA TTTTATCATC ATGAACTAGT GTCTAAATAC CTATGTTTTT TTCAGAAAGC
2521 AAGGAAGTAA ACTCAAACAA AAGTGCGTGT AATTAAATAC TATTAATCAT AGGCAGATAC
2581 TATTTTGTTT ATGTTTTTGT TTTTTTCCTG ATGAAGGCAG AAGAGATGGT GGTCTATTAA
2641 ATATGAATTG AATGGAGGGT CCTAATGCCT TATTTCAAAA CAATTCCTCA GGGGGACCAG
2701 CTTTGGCTTC ATCTTTCTCT TGTGTGGCTT CACATTTAAA CCAGTATCTT TATTGAATTA
2761 GAAACAAGT GGGACATATT TTCCTGAGAG CAGCACAGGA ATCTTCTTCT TGGCAGCTGC
2821 AGTCTGTCAG GATGAGATAT CAGATTAGGT TGGATAGGTG GGGAAATCTG AAGTGGGTAC
2881 ATTTTTTAAA TTTTGCTGTG TGGGTCACAC AAGGTCTACA TTACAAAAGA CAGAATTCAG
2941 GGATGGAAAG GAGAATGAAC AAATGTGGGA GTTCATAGTT TTCCTTGAAT CCAACTTTTA
3001 ATTACCAGAG TAAGTTGCCA AAATGTGATT GTTGAAGTAC AAAAGGAACT ATGAAAACCA
3061 GAACAAATTT TAACAAAAGG ACAACCACAG AGGGATATAG TGAATATCGT ATCATTGTAA
3121 TCAAAGAAGT AAGGAGGTAA GATTGCCACG TGCCTGCTGG TACTGTGATG CATTTCAAGT
3181 GGCAGTTTTA TCACGTTTGA ATCTACCATT CATAGCCAGA TGTGTATCAG ATGTTTCACT
3241 GACAGTTTTT AACAATAAAT TCTTTTCACT GTATTTTATA TCACTTATAA TAAATCGGTG
3301 TATAATTTT
```

FIG. 6Di

```
   1 ATGCAGCGCG TGAACATGAT CATGGCAGAA TCACCAGGCC TCATCACCAT CTGCCTTTTA
  61 GGATATCTAC TCAGTGCTGA ATGTACAGTT TTTCTTGATC ATGAAAACGC CAACAAAATT
 121 CTGAATCGGC CAAAGAGGTA TAATTCAGGT AAATTGGAAG AGTTTGTTCA AGGGAACCTT
 181 GAGAGAGAAT GTATGGAAGA AAAGTGTAGT TTTGAAGAAG CACGAGAAGT TTTTGAAAAC
 241 ACTGAAAGAA CAACTGAATT TTGGAAGCAG TATGTTGATG GAGATCAGTG TGAGTCCAAT
 301 CCATGTTTAA ATGGCGGCAG TTGCAAGGAT GACATTAATT CCTATGAATG TTGGTGTCCC
 361 TTTGGATTTG AAGGAAAGAA CTGTGAATTA GATGTAACAT GTAACATTAA GAATGGCAGA
 421 TGCGAGCAGT TTTGTAAAAA TAGTGCTGAT AACAAGGTGG TTTGCTCCTG TACTGAGGGA
 481 TATCGACTTG CAGAAAACCA GAAGTCCTGT GAACCAGCAG TGCCATTTCC ATGTGGAAGA
 541 GTTTCTGTTT CACAAACTTC TAAGCTCACC CGTGCTGAGA CTGTTTTTCC TGATGTGGAC
 601 TATGTAAATT CTACTGAAGC TGAAACCATT TTGGATAACA TCACTCAAAG CACCCAATCA
 661 TTTAATGACT TCACTCGGGT TGTTGGTGGA GAAGATGCCA AACCAGGTCA ATTCCCTTGG
 721 CAGGTTGTTT TGAATGGTAA AGTTGATGCA TTCTGTGGAG GCTCTATCGT TAATGAAAAA
 781 TGGATTGTAA CTGCTGCCCA CTGTGTTGAA ACTGGTGTTA AAATTACAGT TGTCGCAGGT
 841 GAACATAATA TTGAGGAGAC AGAACATACA GAGCAAAAGC GAAATGTGAT TCGAATTATT
 901 CCTCACCACA ACTACAATGC AGCTATTAAT AAGTACAACC ATGACATTGC CCTTCTGGAA
 961 CTGGACGAAC CCTTAGTGCT AAACAGCTAC GTTACACCTA TTTGCATTGC TGACAAGGAA
1021 TACACGAACA TCTTCCTCAA ATTTGGATCT GGCTATGTAA GTGGCTGGGG AAGAGTCTTC
1081 CACAAAGGGA GATCAGCTTT AGTTCTTCAG TACCTTAGAG TTCCACTTGT TGACCGAGCC
1141 ACATGTCTTC GATCTACAAA GTTCACCATC TATAACAACA TGTTCTGTGC TGGCTTCCAT
1201 GAAGGAGGTA GAGATTCATG TCAAGGAGAT AGTGGGGGAC CCCATGTTAC TGAAGTGGAA
1261 GGGACCAGTT TCTTAACTGG AATTATTAGC TGGGGTGAAG AGTGTGCAAT GAAAGGCAAA
1321 TATGGAATAT ATACCAAGGT ATCCCGGTAT GTCAACTGGA TTAAGGAAAA AACAAAGCTC
1381 ACTTAATGAA AGATGGATTT CCAAGGTTAA TTCATTGGAA TTGAAAATTA ACAGGGCCTC
1441 TCACTAACTA ATCACTTTCC CATCTTTTGT TAGATTTGAA TATATACATT CTATGATCAT
1501 TGCTTTTTCT CTTTACAGGG GAGAATTTCA TATTTTACCT GAGCAAATTG ATTAGAAAAT
```

FIG. 6Dii

```
1561 GGAACCACTA GAGGAATATA ATGTGTTAGG AAATTACAGT CATTTCTAAG GGCCCAGCCC
1621 TTGACAAAAT TGTGAAGTTA AATTCTCCAC TCTGTCCATC AGATACTATG GTTCTCCACT
1681 ATGGCAACTA ACTCACTCAA TTTTCCCTCC TTAGCAGCAT TCCATCTTCC CGATCTTCTT
1741 TGCTTCTCCA ACCAAAACAT CAATGTTTAT TAGTTCTGTA TACAGTACAG GATCTTTGGT
1801 CTACTCTATC ACAAGGCCAG TACCACACTC ATGAAGAAAG AACACAGGAG TAGCTGAGAG
1861 GCTAAAACTC ATCAAAAACA CTACTCCTTT TCCTCTACCC TATTCCTCAA TCTTTTACCT
1921 TTTCCAAATC CCAATCCCCA AATCAGTTTT TCTCTTTCTT ACTCCCTCTC TCCCTTTTAC
1981 CCTCCATGGT CGTTAAAGGA GAGATGGGGA GCATCATTCT GTTATACTTC TGTACACAGT
2041 TATACATGTC TATCAAACCC AGACTTGCTT CCATAGTGGA GACTTGCTTT TCAGAACATA
2101 GGGATGAAGT AAGGTGCCTG AAAAGTTTGG GGGAAAAGTT TCTTTCAGAG AGTTAAGTTA
2161 TTTTATATAT ATAATATATA TATAAAATAT ATAATATACA ATATAAATAT ATAGTGTGTG
2221 TGTGTATGCG TGTGTGTAGA CACACACGCA TACACACATA TAATGGAAGC AATAAGCCAT
2281 TCTAAGAGCT TGTATGGTTA TGGAGGTCTG ACTAGGCATG ATTTCACGAA GGCAAGATTG
2341 GCATATCATT GTAACTAAAA AAGCTGACAT TGACCCAGAC ATATTGTACT CTTTCTAAAA
2401 ATAATAATAA TAATGCTAAC AGAAAGAAGA GAACCGTTCG TTTGCAATCT ACAGCTAGTA
2461 GAGACTTTGA GGAAGAATTC AACAGTGTGT CTTCAGCAGT GTTCAGAGCC AAGCAAGAAG
2521 TTGAAGTTGC CTAGACCAGA GGACATAAGT ATCATGTCTC CTTTAACTAG CATACCCCGA
2581 AGTGGAGAAG GGTGCAGCAG GCTCAAAGGC ATAAGTCATT CCAATCAGCC AACTAAGTTG
2641 TCCTTTTCTG GTTTCGTGTT CACCATGGAA CATTTTGATT ATAGTTAATC CTTCTATCTT
2701 GAATCTTCTA GAGAGTTGCT GACCAACTGA CGTATGTTTC CCTTTGTGAA TTAATAAACT
2761 GGTGTTCTGG TTCAT
```

FIG. 6Ei

```
   1 CTGCAGGGGG GGGGGGGGGG GGGGGCTGTC ATGGCGGCAG GACGGCGAAC TTGCAGTATC
  61 TCCACGACCC GCCCCTACAG GTGCCAGTGC CTCCAGAATG TGGCAGCTCA CAAGCCTCCT
 121 GCTGTTCGTG GCCACCTGGG GAATTTCCGG CACACCAGCT CCTCTTGACT CAGTGTTCTC
 181 CAGCAGCGAG CGTGCCCACC AGGTGCTGCG GATCCGCAAA CGTGCCAACT CCTTCCTGGA
 241 GGAGCTCCGT CACAGCAGCC TGGAGCGGGA GTGCATAGAG GAGATCTGTG ACTTCGAGGA
 301 GGCCAAGGAA ATTTTCCAAA ATGTGGATGA CACACTGGCC TTCTGGTCCA AGCACGTCGA
 361 CGGTGACCAG TGCTTGGTCT TGCCCTTGGA GCACCCGTGC GCCAGCCTGT GCTGCGGGCA
 421 CGGCACGTGC ATCGACGGCA TCGGCAGCTT CAGCTGCGAC TGCCGCAGCG GCTGGGAGGG
 481 CCGCTTCTGC CAGCGCGAGG TGAGCTTCCT CAATTGCTCG CTGGACAACG GCGGCTGCAC
 541 GCATTACTGC CTAGAGGAGG TGGGCTGGCG GCGCTGTAGC TGTGCGCCTG GCTACAAGCT
 601 GGGGGACGAC CTCCTGCAGT GTCACCCCGC AGTGAAGTTC CCTTGTGGGA GGCCCTGGAA
 661 GCGGATGGAG AAGAAGCGCA GTCACCTGAA ACGAGACACA GAAGACCAAG AAGACCAAGT
 721 AGATCCGCGG CTCATTGATG GGAAGATGAC CAGGCGGGGA GACAGCCCCT GGCAGGTGGT
 781 CCTGCTGGAC TCAAAGAAGA AGCTGGCCTG CGGGGCAGTG CTCATCCACC CCTCCTGGGT
 841 GCTGACAGCG GCCCACTGCA TGGATGAGTC CAAGAAGCTC CTTGTCAGGC TTGGAGAGTA
 901 TGACCTGCGG CGCTGGGAGA AGTGGGAGCT GGACCTGGAC ATCAAGGAGG TCTTCGTCCA
 961 CCCCAACTAC AGCAAGAGCA CCACCGACAA TGACATCGCA CTGCTGCACC TGGCCCAGCC
1021 CGCCACCCTC TCGCAGACCA TAGTGCCCAT CTGCCTCCCG GACAGCGGCC TTGCAGAGCG
1081 CGAGCTCAAT CAGGCCGGCC AGGAGACCCT CGTGACGGGC TGGGGCTACC ACAGCAGCCG
1141 AGAGAAGGAG GCCAAGAGAA ACCGCACCTT CGTCCTCAAC TTCATCAAGA TTCCCGTGGT
1201 CCCGCACAAT GAGTGCAGCG AGGTCATGAG CAACATGGTG TCTGAGAACA TGCTGTGTGC
1261 GGGCATCCTC GGGGACCGGC AGGATGCCTG CGAGGGCGAC AGTGGGGGGC CCATGGTCGC
1321 CTCCTTCCAC GGCACCTGGT TCCTGGTGGG CCTGGTGAGC TGGGGTGAGG GCTGTGGGCT
1381 CCTTCACAAC TACGGCGTTT ACACCAAAGT CAGCCGCTAC CTCGACTGGA TCCATGGGCA
1441 CATCAGAGAC AAGGAAGCCC CCCAGAAGAG CTGGGCACCT TAGCGACCCT CCCTGCAGGG
1501 CTGGGCTTTT GCATGGCAAT GGATGGGACA TTAAAGGGAC ATGTAACAAG CACACCGGCC
1561 TGCTGTTCTG TCCTTCCATC CCTCTTTTGG GCTCTTCTGG AGGGAAGTAA CATTTACTGA
```

FIG. 6Eii

```
1621 GCACCTGTTG TATGTCACAT GCCTTATGAA TAGAATCTTA ACTCCTAGAG CAACTCTGTG
1681 GGGTGGGGAG GAGCAGATCC AAGTTTTGCG GGGTCTAAAG CTGTGTGTGT TGAGGGGGAT
1741 ACTCTGTTTA TGAAAAAGAA TAAAAAACAC AACCACGAAA AAAAAAAAAA AAAAAAAAAA
1801 AAAAAAAAAA AAAAAAACCC CCCCCCCGCC CCCCCCCCTG CAG
```

FIG. 6Fi

```
   1 ATGAGGGCTC TGCTGCTCCT GGGGTTCCTG CTGGTGAGCT TGGAGTCAAC ACTTTCGATT
  61 CCACCTTGGG AAGCCCCCAA GGAGCATAAG TACAAAGCTG AAGAGCACAC AGTCGTTCTC
 121 ACTGTCACCG GGGAGCCCTG CCACTTCCCC TTCCAGTACC ACCGGCAGCT GTACCACAAA
 181 TGTACCCACA AGGGCCGGCC AGGCCCTCAG CCCTGGTGTG CTACCACCCC CAACTTTGAT
 241 CAGGACCAGC GATGGGGATA CTGTTTGGAG CCCAAGAAAG TGAAAGACCA CTGCAGCAAA
 301 CACAGCCCCT GCCAGAAAGG AGGGACCTGT GTGAACATGC CAAGCGGCCC CCACTGTCTC
 361 TGTCCACAAC ACCTCACTGG AAACCACTGC CAGAAAGAGA AGTGCTTTGA GCCTCAGCTT
 421 CTCCGGTTTT TCCACAAGAA TGAGATATGG TATAGAACTG AGCAAGCAGC TGTGGCCAGA
 481 TGCCAGTGCA AGGGTCCTGA TGCCCACTGC CAGCGGCTGG CCAGCCAGGC CTGCCGCACC
 541 AACCCGTGCC TCCATGGGGG TCGCTGCCTA GAGGTGGAGG GCCACCGCCT GTGCCACTGC
 601 CCGGTGGGCT ACACCGGACC CTTCTGCGAC GTGGACACCA AGGCAAGCTG CTATGATGGC
 661 CGCGGGCTCA GCTACCGCGG CCTGGCCAGG ACCACGCTCT CGGGTGCGCC CTGTCAGCCG
 721 TGGGCCTCGG AGGCCACCTA CCGGAACGTG ACTGCCGAGC AAGCGCGGAA CTGGGGACTG
 781 GGCGGCCACG CCTTCTGCCG GAACCCGGAC AACGACATCC GCCCGTGGTG CTTCGTGCTG
 841 AACCGCGACC GGCTGAGCTG GGAGTACTGC GACCTGGCAC AGTGCCAGAC CCCAACCCAG
 901 GCGGCGCCTC CGACCCCGGT GTCCCCTAGG CTTCATGTCC CACTCATGCC CGCGCAGCCG
 961 GCACCGCCGA AGCCTCAGCC CACGACCCGG ACCCCGCCTC AGTCCCAGAC CCCGGGAGCC
1021 TTGCCGGCGA AGCGGGAGCA GCCGCCTTCC CTGACCAGGA ACGGCCCACT GAGCTGCGGG
1081 CAGCGGCTCC GCAAGAGTCT GTCTTCGATG ACCCGCGTCG TTGGCGGGCT GGTGGCGCTA
1141 CGCGGGGCGC ACCCCTACAT CGCCGCGCTG TACTGGGGCC ACAGTTTCTG CGCCGGCAGC
1201 CTCATCGCCC CCTGCTGGGT GCTGACGGCC GCTCACTGCC TGCAGGACCG GCCCGCACCC
```

FIG. 6Fii

```
1261 GAGGATCTGA CGGTGGTGCT CGGCCAGGAA CGCCGTAACC ACAGCTGTGA GCCGTGCCAG
1321 ACGTTGGCCG TGCGCTCCTA CCGCTTGCAC GAGGCCTTCT CGCCCGTCAG CTACCAGCAC
1381 GACCTGGCTC TGTTGCGCCT TCAGGAGGAT GCGGACGGCA GCTGCGCGCT CCTGTCGCCT
1441 TACGTTCAGC CGGTGTGCCT GCCAAGCGGC GCCGCGCCGAC CCTCCGAGAC CACGCTCTGC
1501 CAGGTGGCCG GCTGGGGCCA CCAGTTCGAG GGGGCGGAGG AATATGCCAG CTTCCTGCAG
1561 GAGGCGCAGG TACCGTTCCT CTCCCTGGAG CGCTGCTCAG CCCCGGACGT GCACGGATCC
1621 TCCATCCTCC CCGGCATGCT CTGCGCAGGG TTCCTCGAGG GCGGCACCGA TGCGTGCCAG
1681 GGTGATTCCG GAGGCCCGCT GGTGTGTGAG GACCAAGCTG CAGAGCGCCG GCTCACCCTG
1741 CAAGGCATCA TCAGCTGGGG ATCGGGGCTGT GGTGACCGCA ACAAGCCAGG CGTCTACACC
1801 GATGTGGCCT ACTACCTGGC CTGGATCCGG GAGCACACCG TTTCCTGA
```

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/06934

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) :A61K 35/14; C07K 1/00, 14/00; C07H 21/04; A01N 43/04
US CL :435/325, 320.1; 536/23.2, 23.4; 514/44, 822; 530/383, 384, 395
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/325, 320.1; 536/23.2, 23.4; 514/44, 822; 530/383, 384, 395

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

STN (Medline, embase, biosis, caplus), EAST (uspat, epo, derwent), search terms: vitamin k dependent, propeptide, chimer? , gamma carboxylation.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X - Y	WO 88/03926 A1 (NEW ENGLAND MEDICAL CENTER) 02 June 1988 (02-06-88), see entire document, especially p. 24, claims 3-8, p. 25, claims 13 and 16, and pp. 19-22.	1-13 and 18-22 ------- 14-17
Y	WU et al. In vitro Gamma-Carboxylation of a 59-residue Recombinant Peptide Including the Propeptide and the Gamma-Carboxyglutamic Acid Domain of Coagulation Factor IX. J. Biol. Chem. 05 August 1990, Vol. 265, No. 22, pages 13124-13129, especially p. 13125, Fig. 1, p. 13127, Fig. 4, and p. 13128, Table I and Fig. 5.	1-13

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search	Date of mailing of the international search report
08 MAY 2000	30 JUN 2000
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT	Authorized officer Cynthia Laurence Fa [illegible]

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/06934

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y, P	STANLEY et al. Amino Acids Responsible for Reduced Affinities of Vitamin K-Dependent Propeptides for the Carboxylase. Biochemistry. 1999, Vol. 38, No. 47, pages 15681-15687, especially p. 15683, Fig. 1, p. 15685, Table 1, and p. 15686, last 5 lines.	1-22
Y	STANLEY et al. Role of the Propeptide and Gamma-Glutamic Acid Domain of Factor IX for in Vitro Carboxylation by the Vitamin K-Dependent Carboxylase. Biochemistry. 1998, Vol. 37, No. 38, pages 13262-13268, especially page 13263, Col. 2, and p. 13265.	1-13
Y, P	STANLEY et al. The Propeptides of the Vitamin K-dependent Proteins Possess Different Affinities for the Vitamin K-dependent Carboxylase. J. Biol. Chem. 11 June 1999, Vol, 274, No. 24, pages 16940-16944, especially abstract, p. 16942, Fig. 1 and 2, and p. 16944, Col. 1.	1-22